1-1-2003

Metabolic fluxes in developing soybean embryos grown at varying temperatures, quantified by biosynthetically directed fractional 13C-labeling experiments

Vidya Vaancheeswaran Iyer
Iowa State University

Metabolic fluxes in developing soybean embryos grown at varying temperatures, quantified by biosynthetically directed fractional ^{13}C-labeling experiments

by

Vidya Vaancheeswaran Iyer

A thesis submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Major: Chemical Engineering

Program of Study Committee:
Jacqueline V. Shanks (Major Professor)
Mark E. Westgate
Ramon Gonzalez

Iowa State University

Ames, Iowa

2003

Graduate College
Iowa State University

This is to certify that the master's thesis of

Vidya Vaancheeswaran Iyer

has met the thesis requirements of Iowa State University

Signatures have been redacted for privacy

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

NOMENCLATURE

1D	One dimensional
2D	Two dimensional
A-CoA	Acetyl Coenzyme-A
Ala	Alanine
Arg	Arginine
Asx	Aspartate and Asparagine
BLE	Bond Labeling Experiments
F6P	Fructose-6-Phosphate
G6P	Glucose-6-Phosphate
GABA shunt	γ-amino butyrate shunt
Glu	Glutamine
Glx	Glutamine and Glutamate
Gly	Glycine
H6P	Hexose-6-Phosphate
His	Histidine
HPLC	High Performance Liquid Chromatography
HSQC	Heteronuclear Single Quantum Correlation
Ile	Ileucine
Leu	Leucine
Lys	Lysine
Met	Methionine
MFA	Metabolic Flux Analysis
NMR	Nuclear Magnetic Resonance
OAA	Oxaloacetate
P5P	Pentose-5-Phosphate
PEP	Phosphoenolpyruvate
Phe	Phenylalanine
PPP	Pentose Phosphate Pathway
Pro	Proline
Ser	Serine
Suc	Sucrose
T3P	Triose-3-Phosphate
TCA cycle	Tricarboxylic acid cycle

Thr	Threonine
TOCSY	Total Correlation SpectroscopY
Tyr	Tyrosine
Val	Valine

ABSTRACT

Soybeans are used as a major source of raw material for animal feeds as well as other food products. An increase in nutritional value of the seed would render them commercially more attractive. To develop a desired seed product, it therefore becomes necessary that the carbon flow through the metabolic pathways in soybean be clearly understood.

Metabolic flux analysis (MFA) is an effective tool for quantifying intracellular metabolite fluxes in a biological system. This tool can be applied for comparing different phenotypes or genotypes and assessing environmental effects and hence can be extended to suggest further metabolic modifications. The main purpose of this study was to use MFA to assess the temperature effects on central carbon metabolism in developing soybean cotyledons.

Stoichiometric MFA quantifies intracellular fluxes using metabolite balances. As the system becomes more complex, the number of measurements required increases. Plant systems are complex as a result of compartmentation issues, futile cycling and anaplerotic reactions. Our study used a combination of carbon bond labeling experiments and Nuclear Magnetic Resonance (NMR) spectroscopy to provide additional constraints for quantification of the fluxes. Two dimensional (2D) Heteronuclear Single Quantum Correlation Spectroscopy (HSQC) and TOtal Correlation SpectroscopY (TOCSY) analyses were carried out on hydrolyzed protein from the soybean embryos. The resultant NMR spectra were used for evaluating fluxes with the help of a comprehensive mathematical model.

The mathematical analysis involved a single compartment model of soybean metabolism. The reaction network included the primary reactions in metabolism such as glycolysis, pentose phosphate pathway, the tricarboxylic acid (TCA) cycle, anaplerotic pathways, glyoxylate shunt and γ-aminobutyrate (GABA) shunt. The model used the concept of isotopomers to convert the NMR data to metabolic fluxes. The validity of the model was verified by comparing the NMR intensities of experimental and simulated data. The

metabolic flux maps for soybean grown under different temperature treatments have been presented and compared.

In the case of the developing soybean cotyledons, the pentose phosphate to glycolysis flux ratio was different for all the temperature treatments, indicating a 'flexible' node. Similarly analysis of different branchpoints was carried out and the rigidity of the nodes in the network was determined. The anaplerotic reactions were higher for the higher temperature treatments as compared to lower temperature treatments to provide increased amounts of precursor metabolites to compensate for higher amounts of protein produced. The difference in the isotopomer distribution of hydrolyzed sugar molecules and starch sample indicated the likelihood of compartmentation in the upper part of the reaction network. It was concluded that a two-compartment model needs to be constructed to provide more insight on the carbon metabolism in soybean cotyledons.

1. INTRODUCTION

Soybean is used as a major source of edible vegetable oil and high quality protein. The developing soybean seed accumulates around 40 % protein and 21 % lipids of biomass (dry weight). Soy proteins have a huge market as raw material for animal feed as well as other food products. Soy proteins can also be used in industrial non food applications like paper coating, textiles and plastics [1, 2]. Thus, an increased production of protein or oil is highly desired due to the potential industrial applications.

The biomass composition of the seed depends on factors such as the genotype, crop management and environmental conditions (temperature or water) [3, 4]. To assess such effects, a thorough understanding of the metabolism in the system is necessary. Plant metabolism is highly complex with compartmentation, branch points, futile cycling and anaplerotic reactions. Hence, quantification of fluxes can be used as an important tool to resolve these complexities and understand the metabolism in plants. The quantification of the intracellular metabolite fluxes in a biological system is termed metabolic flux analysis (MFA). MFA quantifies fluxes in a network by writing balances around each metabolite with the assumption that the system is in a quasi-steady state. As the system becomes more complex, the metabolite balances have to be complemented with extra cellular measurements such as substrate consumption, secretion of metabolites, biomass composition and intracellular measurements such as labeling information.

The measurement of metabolites, however, can be difficult. Experiments have to be planned taking into consideration the dynamic behavior of the system and measuring the metabolites at different time points. The established methods for the use of extracellular measurements are the gas chromatography mass spectroscopy (GC-MS) and liquid chromatography mass spectroscopy (LC-MS). These techniques efficiently detect sugars, alcohol, organic acids and secondary metabolites [5]. In order to elucidate intracellular branch point distribution, we can use bond labeling experiments. The system under

consideration is fed with a labeled substrate. The distribution of the labeled component throughout the network depends on the metabolic activity of the system. The labeling is reflected in proteinogenic amino acids and other biomass components and can be detected using nuclear magnetic resonance (NMR) spectroscopy [6]. For example, β carbon of alanine and the γ^1 carbon of valine reflect the same carbon atom of their precursor metabolite, pyruvate. Thus, any difference in the intensities of these carbons of alanine and valine can give important information of partitioning of pyruvate into amino acids. NMR is a non destructive technique and hence can also be used to monitor the changes in metabolism of a system *in vivo*.

Thus, a combination of extracellular and intracellular measurements together with the reaction network model can be used for the metabolic flux analysis. Since this network is a highly non-linear system, the analysis is complicated and is computationally expensive. Programs FCAL and 13C-FLUX are used for metabolic flux analysis of central carbon metabolism in *Escherichia coli* and *Corynebacterium glutamicum* systems [7, 8]. We have developed a similar computer program for the metabolic flux analysis in developing soybean cotyledons. A single compartment model was constructed, which included the main reactions of primary metabolism in soybeans such as glycolysis, pentose phosphate pathway (PPP), anaplerotic reactions, TCA cycle, glyoxylate shunt and γ-aminobutyrate (GABA) shunt. Peak fine structures from NMR were translated to metabolic flux information using the isotopomer theory and a mathematical model described below. The combination of external fluxes from the amino acid data and the internal labeling information from the analysis of the NMR spectrum were used to solve the system computationally [9].

A mathematical model representing a reaction network is system specific and hence needs to be tested using experimental data. This model uses two innovative techniques; the singular value decomposition method to obtain a feasible set of fluxes as an initial guess and the Boolean function mapping method to enumerate the isotopomers for the system. The

Boolean function mapping method is efficient and requires minimal input from the user. The mathematical model also incorporates all the known reversible reactions in the system, which is important as the reversibility could affect the flux values to a large extent especially in the case of complex plant systems.

In the case of certain fluxes, the NMR measurement error can translate to a very high confidence interval in the estimation of the flux, leading to identifiability problems [8]. The reaction network is represented with a mathematical model supplemented with external measurements. The information about different parts of the pathway is obtained from measurements pertaining to that particular part of the pathway. For example, flux analysis of the reversibilities of the transketolase and transaldolase reactions in the pentose phosphate pathway depend primarily on the labeling information obtained from the PEP family of the amino acids and His. If the external measurements or the independent NMR information are not sufficient to estimate the fluxes or if the error in the measurements is large, then the fluxes become 'structurally unidentifiable'. This problem can be solved by increasing the number of measurements or providing the system with low noise level measurements. However, in some cases, the relationship between the NMR measurements and the fluxes are very complicated. In such cases, the fluxes become 'statistically unidentifiable' and a very low noise level can translate into large confidence intervals for the fluxes. Thus, in the case of a statistically unidentifiable flux, the model cannot estimate the flux irrespective of redundant measurements pertaining to that flux. The independent reactions of the network also need to be identified and NMR measurements giving information pertaining to that reaction become a mandatory requirement for solving the system. Such issues were studied in detail in the course of developing the model for the soybean system. Organic acids effluxes and substrate consumption were found to be some of the important parameters for a more accurate determination of the metabolic fluxes.

The main objective of this study was to assess the effect of varying temperature treatments on metabolic fluxes in developing soybean embryos using labeled sucrose as the substrate. The soybean cotyledons were cultured *in vitro* with 10 % uniformly labeled ^{13}C sucrose under different temperature treatments. The biomass was extracted and separated into oil, protein and starch. This work was carried out at Dr. Westgate's lab in the agronomy department at Iowa State University. The proteinogenic amino acids obtained from hydrolysis were quantified and used as external inputs for the system. These amino acids reflect the ^{13}C labeling patterns of their precursor metabolites directly and provide important information on the intracellular fluxes. Branch point identification was achieved by NMR, which provides additional constraints and gives information about the flow of carbon label through the reaction network. 2D Heteronuclear Single Quantum Correlation Spectroscopy (HSQC) and TOtal Correlation SpectroscopY (TOCSY) analyses were carried out on the hydrolyzed protein samples. The HSQC analysis detects the distribution of the labeling through the reaction network. Also, since we have unlabeled glutamine as a carbon source in addition to the labeled sucrose, there is a dilution of ^{13}C in the system. The 2D TOCSY analysis provides the enrichment information of each carbon atom of amino acids. The single compartmental model was used to develop metabolic flux maps for five temperature treatments spanning below optimal and above optimal conditions for protein accumulation. The mathematical model and the data presented can be used to optimize growth conditions for obtaining the desired biomass composition. This model, in future also could be extended to direct metabolic engineering of soybean lines to partition more carbon into protein or oil.

2. LITERATURE REVIEW

2.1 Quantification of fluxes in plants

To produce plants with desired chemical properties, it is important to understand plant metabolism in detail since genetic manipulations have seldom produced the desired results [10]. Many techniques such as gas chromatography-mass spectroscopy [11], liquid chromatography-mass spectroscopy (LC-MS) and NMR have been developed to measure metabolites, to give information about central carbon metabolism [12]. The metabolite data can be used for the quantification of fluxes in the reaction network and are helpful in understanding the key points of regulation in the pathways of central metabolism [5]. Once the reaction network of a plant system is modeled, genetic manipulations or comparison between phenotypes can be studied in detail to yield products with desired properties. Thus, quantification of fluxes is an important tool, which needs to be explored more towards understanding plant metabolism.

Although, it is an established technique, NMR has not been used to study plant metabolism to a large extent [13]. The amount of information from NMR analysis, however, is enormous. For example, β carbon of alanine and the γ^1 carbon of valine reflect the same carbon atom of pyruvate from which they originate. Thus, any difference in the intensities of these carbons of alanine and valine can give important information of partitioning of their precursor metabolite pyruvate into corresponding amino acids. Such information is particularly useful in the case of plant systems where compartmentation possibility is high [12, 13]. Bond labeling experiments have been found suitable for the quantification of fluxes in complex eukaryotic systems like plants [14]. One of the early applications included the quantification of metabolic fluxes in maize root tips. 1D ^{13}C and ^{1}H experiments were carried out to analyze the enrichment data [15]. Other applications using retrobiosynthetic analysis included quantifying biosynthesis of anthraquinones in *Rubia tinctorum* using a combination

of 1D and 2D NMR experiments [16] and amino acid and starch biosynthesis in developing maize kernels [17, 18]. An *in situ* NMR experiment has also been carried out on *Catharantus roseus* to study the changes in the primary metabolism as a function of time [19].

2.2 Metabolic flux analysis

Metabolic flux analysis (MFA) involves the quantification of intracellular steady state fluxes in the cell, using metabolite balances and extra cellular metabolite measurements. It is an important tool, which helps in comparing phenotypes or genetic variants as well as suggesting possible targets for genetic modification [6, 19, 20]. MFA requires that each metabolite in the entire reaction network under investigation be represented by atleast one reaction. The reactions are listed and a balance around each metabolite is written to create a stoichiometric model [21, 22]. Thus in vector notation,

$$d\mathbf{X}/dt = \mathbf{r} - \mu\mathbf{X}$$

where, **X** represents the metabolite under consideration, **r,** the rate of formation of the metabolite and μ is the biomass growth rate. The assumption is a pseudo-steady state *i.e.* the concentrations of the intracellular metabolites do not change with time. Hence, we have,

$$\mathbf{r} - \mu\mathbf{X} = 0$$

The dilution due to the biomass growth is generally small and the second term can be neglected and we have,

$$\mathbf{r} = \mathbf{G}^{\mathrm{T}}.\mathbf{v} = 0$$

where, **v** is the vector containing the fluxes and **G** is the stoichiometric matrix. The measured and calculated fluxes can be partitioned into $\mathbf{v}_m$ and $\mathbf{v}_c$ respectively. Correspondingly, the stoichiometric matrix can be partitioned into $\mathbf{G}_m$ and $\mathbf{G}_c$. Thus, knowing $\mathbf{v}_m$ and **G**, we can calculate $\mathbf{v}_c$, the set of unmeasured intracellular fluxes. If there are J reactions and K metabolites then the degrees of freedom F, is given by,

$$F = J - K$$

If the number of extracellular measurements are same as F it is an exactly determined system, if greater than F, an overdetermined system and if less than F, an underdetermined system. In the case of an overdetermined system the extra measurements can be used to check the validity of the model. To solve an underdetermined system one needs to resort to cofactor dependent balances (NADPH/ NADH) [22, 23]. However, the NADPH, NADH and ATP balances are not closed in reality due to futile cycles and incomplete pathway knowledge. Stoichiometric MFA also fails in certain cases of parallel pathways and metabolic cycles [24]. Thus, it becomes necessary to provide additional information and also elucidate flux distribution at branch points. For larger systems, flux analysis becomes more difficult as the number of measurements required increases.

2.3 Branchpoint Information by Nuclear Magnetic Resonance Spectroscopy

Flux distributions at branchpoints can be obtained from ^{13}C labeling experiments. These experiments involve feeding the biological system ^{13}C labeled substrates such as glucose or sucrose. The experiment results in the redistribution of the label throughout the network by the substrate being broken down into corresponding metabolites. The extent of the enrichment of the metabolites can be measured by NMR [6, 25] or Mass Spectroscopy (MS) [11]. NMR and MS also can be used in combination to provide additional information [26].

NMR is a non-intrusive technique, which can be used to study the changes in the metabolic state of a system *in vivo*. 2D NMR can efficiently and unambiguously be used to assign peaks of different metabolites in a single experiment. In organisms and plant systems, amino acids are known to reflect their precursor metabolites directly. Thus, NMR can give us important information on the central carbon metabolism [12, 13]. In an NMR measurement, the spin of the ^{13}C nucleus is detected. The variation in the frequency of resonance is called the chemical shift and is characteristic for each nucleus. A 1-dimensional (1D) proton

spectrum gives the enrichments of all the protonated carbon present in the sample [25]. The enrichment information can be used to elucidate the metabolic flow in the pathway. The information in a 1D spectrum is in the form of a number of peaks, which may overlap and make the analysis difficult. However, 2-dimensional (2D) NMR is more convenient as the peak is dispersed in 2 dimensions and there is less ambiguity in the assignment of each nucleus.

Szyperski [15] developed a biosynthetically directed fractional labeling or bond labeling experiment (BLE) [27], in which a mixture of uniformly labeled and unlabeled substrate was used for an *in vitro* culture of *Escherichia coli*. The breakdown of the labeled substrate into subsequent reactions was detected using a 2D [^{1}H, ^{13}C] COSY (COrrelation SpectroscopY) experiment. 2D COSY detects the interaction of protons directly attached to carbons. Since the detection is restricted to large one bond scalar coupling, only adjacent carbon atoms are detected and information about the origin of the carbon atoms of a metabolite is obtained. However, isotopic enrichment information is not obtained from BLE's as it is uniform for all the carbon atoms of the metabolites [6]. A 2D HSQC (Heteronuclear Single Quantum Correlation spectroscopy) analysis also can be carried out to detect the ^{13}C-^{13}C scalar couplings [12]. In the HSQC experiment the magnetization is transferred from the proton to the carbon and then back to the proton. The HSQC experiment proves advantageous even when the coupling between the carbon atoms is small, which helps in representing the relative abundances of the isotopomers. These abundances can be translated to flux information, which is explained in the next section[26, 28].

Other experiments include using selectively labeled substrate to obtain information about central metabolism [25]. However, fractional labeling information which is obtained from selectively labeled substrates yield only n measurements as opposed to 2^n measurements from BLE. The fractional labeling information can be used as additional constraints for the stoichiometric model.

2.4 Isotopomer and Bondomer theory

2.4.1 Isotopomer model

For a metabolite with n carbons, there are 2^n labeling patterns possible. These isotope isomers which indicate different labeling patterns are defined as isotopomers. The abundances of the 2^n isotopomers depends on the extent of labeled substrate (for example, glucose or sucrose), the natural abundance of the unlabeled substrate and the metabolism [12]. For a 3 carbon molecule, there are $2^3 = 8$ possible isotopomers. Figure 2.4.1 explains different multiplet patterns that can be obtained in a 2D NMR experiment from a 3 carbon fragment when the central carbon atom is being observed. These peak patterns can be compared to the isotopomer abundances simulated by constructing and solving Isotope Mapping Matrices (IMM). IMM uses the concept of isotopomer distribution vectors (IDV) and reaction stoichiometry [28]. IMM are analogous to the Atom Mapping Matrices (AMM) used to calculate the TCA flux ratios in a hybridoma cell line [29]. The concept of IMM was used successfully for metabolic flux analysis of *Escherichia coli* [30]. Klapa *et al.* (1999) used the concept of IMM by considering only the realizable isotopomers depending on the biochemistry. They used selectively labeled substrate in this case as compared to a mixture of uniformly labeled and unlabeled substrate in the case of Schmidt *et al* (1999).

2.4.2 Bondomer model

Recently, two groups introduced a new concept of bondomers, which are similar to isotopomers, except that the bonds instead of the carbon atoms are being followed [31, 32]. Bondomers are molecules of the same metabolite, which have different bond integrities for different carbon-carbon bonds (Sriram and Shanks, 2003). The bond integrity of a covalent bond indicates whether or not the two carbon atoms under consideration arise from the same substrate. If they arise from the same substrate the bond integrity is 1, else it is 0. A metabolite with n carbons is represented by the following equation and b_i indicates the bond

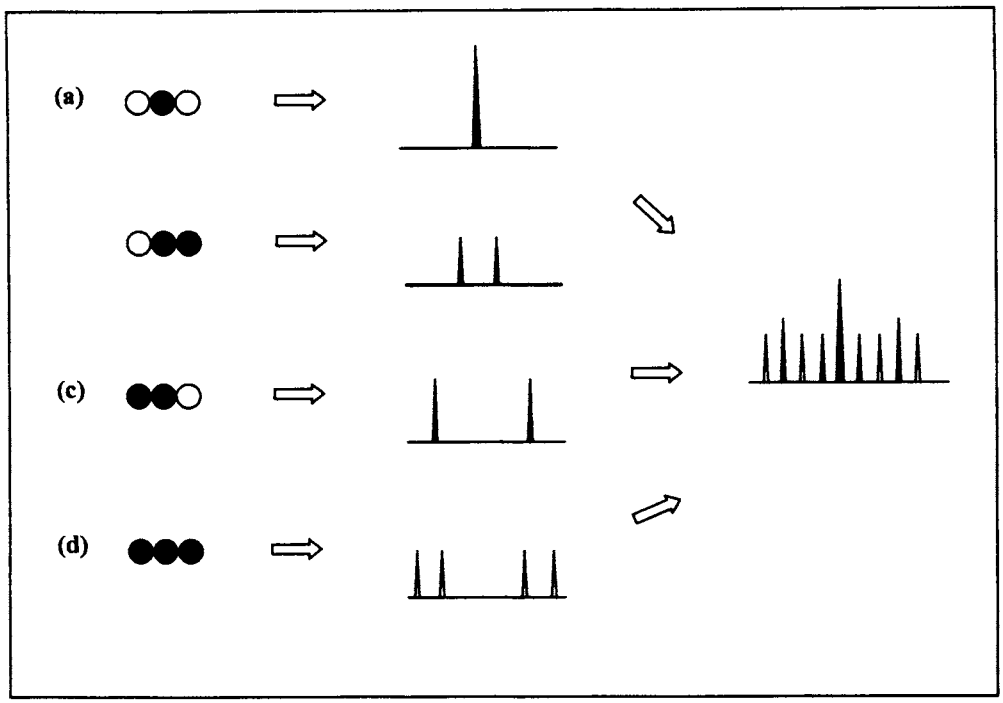


Figure 2.4.1 Multiplet peak patterns of a 3 carbon metabolite using NMR. When observing the central carbon (that is ^{13}C), patterns observed are: (a) singlet when the carbon is attached to two ^{12}C atoms (b) doublet when the carbon is attached to one ^{12}C and one ^{13}C atom (c) another doublet when the carbon is attached to one ^{13}C and one ^{12}C atom (d) double doublet when the carbon is attached to two ^{13}C atoms [24].

integrity of the i^{th} bond.

$$\begin{matrix} & b_1 & b_2 & b_3 & b_4 & & & b_{n-1} \\ 1 & - & 2 & - & 3 & - & 4 - 5 \ldots\ldots\ldots\ldots - & n \end{matrix}$$

The bondomer number for the bondomers of the molecule is calculated as

$$k = \sum_{i=0}^{n-1} 2^{i-1} b_i + 1$$

For example, consider a 3 carbon molecule 1-2-3. If bond integrity of 1-2 is 0 and 2--3 is 1 then the bond number for the bondomers of the molecule is calculated as

$$0 * 2^0 + 1 * 2^1 + 1 = 3$$

Hence, the combinations possible are tabulated in Table 2.4.1. Thus, for an n carbon molecule there are 2^{n-1} possible bondomers. Thus, the application of bondomers theory to metabolic flux analysis is very clearly computationally less expensive than isotopomers [32].

Table 2.4.1 Bondomer abundances possible in a 3 carbon metabolite

Metabolites	1-2	2-3	Bondomer #
1-2-3	0	0	M_1
1--2-3	1	0	M_2
1-2--3	0	1	M_3
1--2--3	1	1	M_4

2.5 Mathematical Modeling and applications

The first step in metabolite flux balancing (MFB) involves writing down all the stoichiometric reactions, which form the biochemical network under consideration. Steady state balances are written around each metabolite in the network. External measurements (for example secreted metabolites, substrate consumption etc) are required, which provide constraints to the system. The external measurements can be used as inputs to solve the biochemical network. In the case of an underdetermined system, the network can be solved by a combination of the external measurements and optimizing the system for a particular objective like biomass growth, using a platform like MATLAB [21]. The problem of underdetermined systems was solved when NMR and MS data could be used to provide additional constraints to the system. The tricarboxylic acid (TCA) cycle flux ratios in a hybridoma cell were calculated using a combination of NMR data, atom mapping matrices (AMM) and the Gauss-Siedel iterative method [29]. Many of these models, however, considered only unidirectional reactions in the pathway, which leads to an error in the flux

estimations. The effect of reversibility was later incorporated in a mathematical model by considering the reactions in the pentose phosphate pathway [27, 33]. Wiechert *et al.* (1997b) extended this to the primary metabolism reactions in glycolysis, pentose phosphate pathway, TCA cycle and anaplerotic reactions in *Corynebacterium glutamicum.*

In BLE's, proteinogenic amino acids, which can be obtained from hydrolysis of the biomass protein, provide additional constraints to solve the network. These amino acids reflect their precursor metabolite directly and hence there is no need to break them down further and are well spread across the pathways of primary metabolism thereby providing constraints for different parts of the reaction network [6]. BLE also introduced the concept of bond connectivity. A bond is said to be intact if the 2 adjacent carbons in a metabolite originate from the same molecule. A system of probabilistic equations was proposed, which provides insight to the bond connectivity. The probabilistic theory was used in conjunction with metabolic flux ratio analysis (METAFoR) to elucidate a metabolic network. A program called FCAL (Flux ratios from Correlated Amino acid ^{13}C Labeling patterns) was introduced, which calculates flux ratios at different points in the network that was exemplified in a halophilic archaeon *Haloarcula hispanica* [7].

Simultaneously, software was developed for metabolic flux analysis with a Monte-Carlo statistical analysis and exemplified for flux analysis in *Aspergillus niger* strain [34]. Here it was inferred that fluxes were estimated best using an evolutionary algorithm for error minimization. The Monte-Carlo procedure was used to identify the statistical quality of the estimated fluxes. A mathematical model was also developed, which included only the physically realizable isotopomers and was solved in a non-iterative manner for systems of *Escherichia coli, Pseudomonas* AM1 and mammals [26, 35]. However, the physically realizable isotopomer method applies only to specific systems and specific kind of labeling experiments. 13C-FLUX, a universal software framework was developed for metabolic flux analysis [8]. A comprehensive explanation about the experimental design and working of the

model is exemplified on a *Corynebacterium glutamicum* system for the determination of the anaplerotic fluxes [8]. The concept of cumomers (cumulative isotopomers) was introduced to analytically solve isotopomer balances [36, 37]. The concept of cumomers was extended to bondomers, which led to the introduction of cumulative bondomers and also allowed an analytical solution of the bondomers balances. The method was shown to be computationally more efficient as compared to cumulative isotopomers [32]. The concept of bondomers and cumulative bondomers has been demonstrated to be efficient and numerical innovations have been reported for metabolic flux analysis [9]. A computer program also has been developed to apply the concept of cumomers for a two substrate model for soybean cotyledons [38]. Other applications in the field of flux analysis included monitoring metabolic flux in lysine synthesis in *Corynebacterium glutamicum* using 1D NMR [39], flux analysis in *Bacillus subtilis* grown on two carbon substrate mixtures [40, 41] and flux response to pyruvate knockout in *Escherichia coli* [42].

2.6 Temperature stress in soybean

In our present study, the system under consideration is the developing soybean cotyledon. Soybean is a very good source of edible oil and protein, which have many industrial applications [1, 2]. The mass and composition of soybean seeds change with varying reproductive growth stages [43]. The reproductive growth stage R5, R6 and R7 characterize beginning seed, full seed and beginning maturity respectively. Dornbos *et al.* (1986) showed that the majority of the dry weight, protein, oil and total sugar accumulated between R5 and R7 respectively. The moisture content decreased between R5 and R7 steadily. In environmental studies in Argentina, soybean was grown under different conditions of latitude, longitude, altitude and temperature. There was no significant correlation between the genotypes and environmental conditions. It was concluded that the selection of a genotype from their time of maturity was a more important factor [3]. Work

has also been carried out regarding pinitol accumulation in soybean plants in response to temperature stress. The pinitol content increased in leaves and stems under higher temperature conditions, whereas the lower temperatures caused little change in the pinitol content [44]. Thus, temperature, growth stage and environment are important factors contributing to the final seed composition.

Studies have been carried out to understand the effect of varying carbon and nitrogen concentrations on growth of soybean embryos *in vitro*. A decreased glutamine concentration affected both the protein and oil content whereas an increase in the glutamine concentration increased the protein content but not the oil content. An increase in sucrose concentration increased the dry weight and protein content but the protein content was lower at the maximum sucrose concentration of 150 mM because of enhanced dry weight [45]. In another parallel study, it was observed that a minimal concentration of 17 mM of glutamine was sufficient for dry weight accumulation without the need for nitrogen accumulation [4]. The concentrations of sucrose (146mM) and glutamine (37mM) for the *in vitro* culture system of soybean cotyledons in our study were determined from these studies [4, 46].

The goal of the present study is to investigate the effects of varying temperature on the central carbon metabolism in developing soybean embryos (Variety *Evans*). The amount of biomass accumulated is representative of the amount of carbon taken into the system. The labeled substrate sucrose (and the unlabeled substrate) are metabolized through the network into amino acids and ultimately accumulated in the protein. Thus, a retrobiosynthetic approach using BLE's and NMR spectroscopy were applied for the metabolic flux analysis of the soybean embryo. The comparison of key regulatory points will help in analyzing the differences in carbon flux through the metabolic pathway due to varying temperature conditions.

The following chapter gives the detailed protocol followed for the *in vitro* culture, biomass component extractions, protein hydrolysis and amino acid analysis, NMR measurements and the mathematical flux model.

3. MATERIALS AND METHODS

3.1 Soybean culture system and temperature treatments

Soybean (*Glycine max* L. Merr.), variety 'Evans' was grown in 20 L pots filled with soil mix in the growth chamber at a day/night temperature of 27°C/ 20°C under 14 h days with a photosynthetic photon flux density of 650 μE m^2/s at canopy level from a combination of cool white, VHO fluorescent Philips F 96 T12 (Somerset, NJ, USA) and 100 W incandescent lamps. Pots were over seeded and then thinned to three plants per pot at growth stage V1 [4]. The plants were fertilized with 0.4-0.5g/pot (20-20-20) of Peters professional fertilizer (Peter Fertilizer Products, W.R.Grace & Co., Fogelsville, PA) every week. Plants were irrigated with tap water every day. Plants were grown at 27°C/ 20°C until 10 days after flowering (DAF). Air temperature was adjusted to 35°C/ 27°C (high, H), 27°C/ 20°C (medium, M), or 20°C / 12°C (low, L) during seed filling. Seeds were selected for *in vitro* culture beginning 18 DAF at 100 to 120 mg fresh weight per seed. Table 3.1.1 summarizes the temperature treatments under which the cotyledons were grown *in vitro*.

The procedure for *in vitro* culture was modified from Obendorf et al (1983). The embryos were excised from seeds taken from pods borne on the central nodes. After seed coats were removed, cotyledons were surface sterilized with 20% regal liquid bleach (sodium hypochlorite 5%) for 15 min, transferred to 70% ethanol for 20 min, and then rinsed three times in sterile distilled water. One cotyledon was placed in a 50 ml Erlermeyer flask containing 4 mL of sterile medium containing 146 mM sucrose (10% ^{13}C sucrose), 37 mM glutamine, vitamins, and micronutrients at pH 5.5. Each temperature treatment had triplicate samples. The medium was replaced every 3 days. Cotyledons were harvested after 6 days, rinsed with sterile medium and frozen in liquid nitrogen and lyophilized. Freeze dried cotyledon tissue was used for oil, protein, starch and sucrose analysis.

Oil was extracted from approximately 100 mg lyophilized embryo tissue in hexane and weighed. Protein in the defatted pellet was separated from non- protein nitrogen by precipitation in approximately 20 % trichloroacetic acid (TCA) and quantified by micro Kjeldahl procedures. Starch and sucrose in the defatted powder were quantified by glucose oxidase after hydrolysis. The extraction protocols are explained in detail in Appendix A. These experiments were carried out in Dr. Westgate's laboratory in Agronomy Department at Iowa State University.

Table 3.1.1 Temperature treatments for *in vivo* and *in vitro* soybean culture

Number of replicates	Sample #	Growth chamber temperature, °C (*in vivo*)	Incubator chamber temperature, °C (*in vitro*)	Growth/Incubator chamber
3	1/2/3	35 ("high, H")	35 ("high, H")	H/H
3	4/5/6	27 ("medium, M")	35 ("high, H")	M/H
3	7/8/9	27 ("medium, M")	27 ("medium, M")	M/M
3	10/11/12	27 ("medium, M")	20 ("low, L")	M/L
3	13/14/15	20 ("low, L")	20 ("low, L")	L/L

3.2 Protein hydrolysis and amino acid analysis

The protein extracted from cotyledons cultured under different temperature treatments was hydrolyzed to recover the amino acids from the sample. The optimized procedure for protein hydrolysis is as follows. The protein was placed in a vacuum hydrolysis tube (Pierce chemical company, product # 29564) and hydrochloric acid (constant boiling point, 6N, Pierce chemical company, code # 0024350 1803580) was added. The tube was then sealed and subjected to vacuum. It was flushed with N_2 to remove any traces of air

left. This was repeated twice to ensure complete vacuum and remove all traces of O_2 from the tube. The tube was then placed in a heating block for around 4 hours at 140°C. The sample was removed from the heating block and was vaporized in Rapidvap (Labconco, model # 7900002) for the removal of HCl in two stages. The first stage consists of a 2-hour cycle at 45°C and a pressure of 20 bars. The rotor speed was set to 35 % of maximum. The first stage ensures the vaporization of HCl. In the second stage, the sample was reconstituted in 2 ml of de-ionized water. The second stage was carried out at atmospheric pressure and room temperature at 80-90 % of maximum speed for 20 minutes, ensuring that the sample is completely reconstituted in the water.

The final step was freeze-drying. The rapid vaporization stage is important, as the freeze dryer cannot handle HCl vapors. The sample was freeze dried in freeze dry system (Labconco, model: Freezone 4.5) for 78 hrs at -35°C and 133 mTorr pressure. The dried protein sample was then reconstituted in 250-400 µl of deuterium oxide (D_2O, Sigma-Aldrich, product # 45336-6) and vortexed thoroughly. The hydrolyzed protein sample was analyzed by high performance liquid chromatography (HPLC) to determine the mole percentages of the amino acids. The HPLC analysis was carried out at the HPLC protein facility (Biochemistry, Biophysics and Molecular Biology department) at Iowa State University). The amount of hydrolyzed protein sample required for amino acid analysis was 10 µl.

3.3 NMR sample preparation and analysis

3.3.1 NMR sample preparation

The sample was prepared in NMR tubes (Wilmad 528-PP). The sample preparation consists of the following steps

1. The hydrolyzed protein sample was reconstituted in approximately 250-400 µl of D_2O after freeze-drying.
2. 10 µl of the standard 2, 2-dimethylsilapentane-5-sulfonic acid (DSS, Sigma-Aldrich, product # 17883-7) was then added.
3. The pH of the sample was then adjusted between 0.5-1.0 by adding DCl (deuterium chloride, Sigma-Aldrich, product # 17672-9).

The total sample volume required to avoid inhomogenity in the magnetic field was around 500 µl or more.

3.3.2 Processing of 2-dimensional NMR spectra

The ^{13}C-labeling pattern, which is reflected in the proteinogenic amino acids, can be analyzed using NMR spectroscopy. We used two analytical approaches, the HSQC (^{1}H, ^{13}C Heteronuclear Single Quantum Correlation spectroscopy) and the TOCSY (TOtal Correlation SpectroscopY). The HSQC analysis determines the degree of coupling between the adjacent carbon atoms. Also, since we have unlabeled glutamine as a carbon source in addition to the labeled sucrose, there is a dilution of ^{13}C in the system. The 2D TOCSY analysis provides the enrichment information of each carbon atom of amino acids.

The 2D spectra were collected on a Bruker advance DRX 500 MHz spectrometer in the NMR facility (Biochemistry, Biophysics and Molecular Biology department) at Iowa State University). The NMR spectra were acquired and processed using the Xwinnmr (Bruker) software. The reference zero ppm was set using the methyl signal of DSS as an internal standard. The resonance frequency of ^{13}C and ^{1}H were set to 125.7 MHz and 499.9 MHz, respectively. The sample temperature was maintained at 25°C. Table 3.3.1 lists the NMR parameters that were used during the experiments. The number of scans ranged from 16-32 and was inversely proportional to the concentration of the sample. The spectral width of 5028.05 Hz in the ^{13}C dimension corresponded to 40 ppm. The experiments were set up so

that peaks outside this sweep width would appear at a chemical shift of δ-40. To elucidate, the alpha peaks appear in the range of 54-62 ppm. In our spectrum it was folded over and appears at 14-22 ppm. The window function used for processing was a sine function with a 90° phase shift. After zero filling and processing, there were 1024 complex points in both dimensions.

Table 3.3.1 Parameters used for NMR analyses

HSQC		
Parameters	^{1}H	^{13}C
Spectral width	5482.26 Hz	5028.05 Hz
Complex datapoints	1024	900
TOCSY		
Parameters	*Values*	
Power level	5.6 KHz	
Isotropic mixing sequence	DIPSI-2 (76 msec.)	

The HSQC and TOCSY spectra were analyzed using the free software NMRview (version 4 and 5 available at http://www.nmrview.com/). The software enabled extraction of 1D slice of the peaks and quantifying the peak intensities.

3.4 Mathematical modeling of the reaction network

The Shanks lab has developed computer software for the mathematical modeling of the reaction network [9]. The model includes a singular value decomposition method to guess an initial set of fluxes. These fluxes are converted to isotopomer distributions using a Boolean function mapping method. The simulated and experimental (from NMR data) distributions are compared and the error between them is minimized using simulated annealing. The fluxes are refined using the Monte Carlo method and the process is repeated

till the global optimum is found. Once the loop converges for the guessed set of initial fluxes, the program estimates a new set of initial fluxes and multiple simulations can be carried out. The package also includes statistical analysis of the results and provides 90% confidence intervals for the fluxes and reversibilities (Figure 3.4.1). The variability arises because each NMR measurement has an error due to the noise level in the spectrum. The program assumes a normal distribution of the error around the NMR measurement. Every time a new simulation is carried out, the program guesses a new set of initial fluxes and compares the simulated intensities with the randomly perturbed experimental NMR intensities.

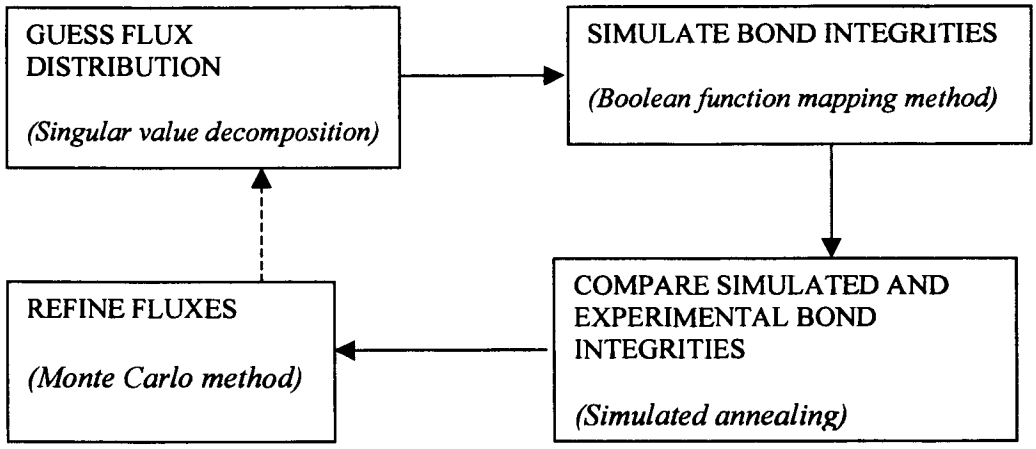


Figure 3.4.1 Algorithm used for simulation of intracellular fluxes

Sometimes, peaks may overlap as a result of which an observed peak is actually a combination of more than one peak. The overlap makes deconvolution of the peaks necessary. The Shanks lab has developed a computer program for this purpose. The program uses the technique of simulated annealing to minimize the error between the simulated and experimental data. The reason for the overlap of the peaks is explained more in detail in Chapter 4.

4. RESULTS AND DISCUSSION

The results are presented in three parts. Initially, the soybean culture experiments and protein quantification are discussed. This is followed by the single compartment model suggested for the primary metabolism in soybean. The calculation of external fluxes and analysis of the NMR data are discussed in detail. Finally, the computational results for the five temperature treatments are presented. The validity of the model used for flux analysis is discussed by comparing the simulated and experimental intensities. In addition, a comprehensive comparison of the fluxes of all the temperature treatments is presented.

4.1 Extraction of Protein, Starch and Lipids from soybean embryos

As discussed in Chapter 3, the soybean cotyledons were grown *in vitro* under five different conditions to study the effect of temperature on partitioning of sucrose and glutamine into biomass. The soybean cotyledons were transferred 18 days after flowering (DAF) from the growth chamber to the *in vitro* culture in the incubator chamber for a period of 6 days. The cotyledons were then removed from the *in vitro* culture for extraction of oil, protein and starch from the biomass (procedures in Appendix A). The results are shown in Figure 4.1.1 and were obtained from experiments carried out in Dr. Westgate's laboratory in Agronomy Department at Iowa State University.

Figure 4.1.1 shows that the general trend for dry weights per cotyledon accumulated over a period of 6 days decreases with decreasing temperature conditions. It was also observed that dry weight for the medium/low (M/L) treatment was the same as the low/low (L/L) treatment. The percentage of protein was highest for the medium/medium (M/M) treatment. The starch increased with decreasing temperature, whereas the oil showed the opposite trend. From the biomass dry weight and the protein percentages, it is possible to calculate the total amount of protein in the cotyledon.

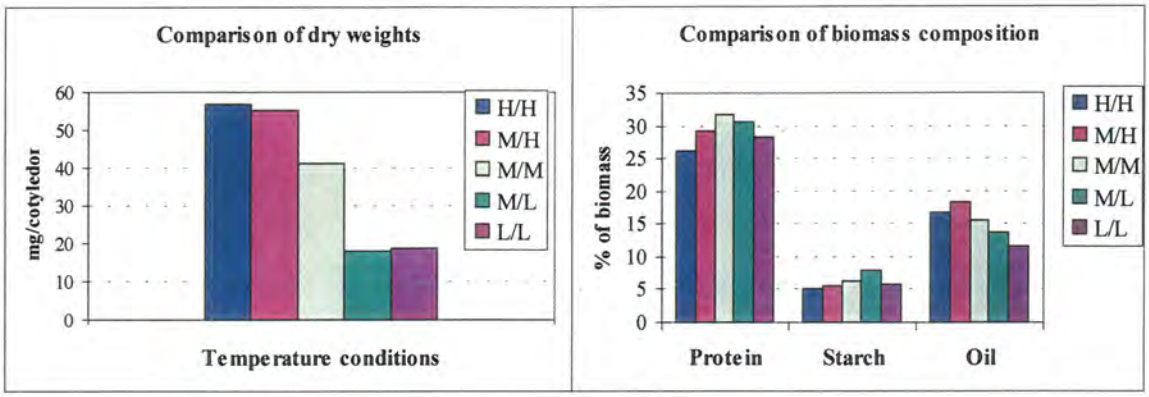


Figure 4.1.1 Final biomass and composition for soybean cotyledons under different temperature treatments (initial cotyledon dry weight was 9.6 mg with a standard deviation of 0.6)

4.2 Amino acid quantification

The protein was measured using the micro Kjeldahl procedure in Dr. Westgate's lab in Agronomy Department at Iowa State University (Appendix A). The amount of protein in mg was estimated from the biomass accumulated over 6 days and the measured protein percentages. Table 4.2.1 shows the amounts of protein under different temperature treatments. The observed trend showed that the protein quantities were in comparatively higher range for the high/high (H/H), medium/high (M/H) and medium/medium (M/M) temperature treatments. Although the percentage of protein was the highest in M/M (optimum condition for protein accumulation), the total protein was highest for the M/H, followed by the H/H and then the M/M temperature treatments. The higher values of protein were obviously because the dry weights were higher for the H/H and M/H as compared to the M/M temperature treatment. At the lower temperature treatments, the protein quantities were affected due to a decrease in the accumulation of biomass.

Each protein sample was hydrolyzed into the corresponding amino acids. The hydrolysis procedure was followed by an HPLC analysis, which was carried out at the

Table 4.2.1 Comparison of protein accumulation for the five temperature treatments

Temperature Treatments	Sample # (*)	Biomass (mg)	Protein % of Biomass	Protein measured (mg)
H/H	1/2/3	56.8	26.3	15.0
M/H	5/6	55.2	29.3	16.2
M/M	7/8	41.3	31.7	13.1
M/L	11	17.9	30.7	5.5
L/L	15	18.9	28.2	5.3

* indicates samples pooled for analysis

protein facility in BBMB Department at Iowa State University. Figure 4.2.1 shows the comparison of the HPLC analysis of the hydrolyzed protein for the different temperature treatments. To test the HPLC technique, BSA (bovine serum albumin, a protein standard) was hydrolyzed for 2 hours and 3 hours and their HPLC analysis was compared with literature values of BSA composition. The deviations between the results of the two hydrolysis treatments and the published values were in the range of 0-5 % as shown in Figure 4.2.2. These results confirmed the accuracy of the HPLC analysis, which was used to compare the hydrolyzed protein from the soybean cotyledons. The results from the five temperature treatments were compared using the T-test. The statistical difference of the remaining temperature treatments from the M/M temperature treatment was evaluated and are listed in Table B.1 in Appendix B. Amino acids like Thr, His, Ala and Leu were significantly different from the M/M treatment. Methionine, in extremely low quantities, also showed a significant difference. There was no particular trend observable in our study for amino acid composition for different temperature treatments. The exception was for amino acids like Leu, His and Met, where there was a statistical difference. The amino acids

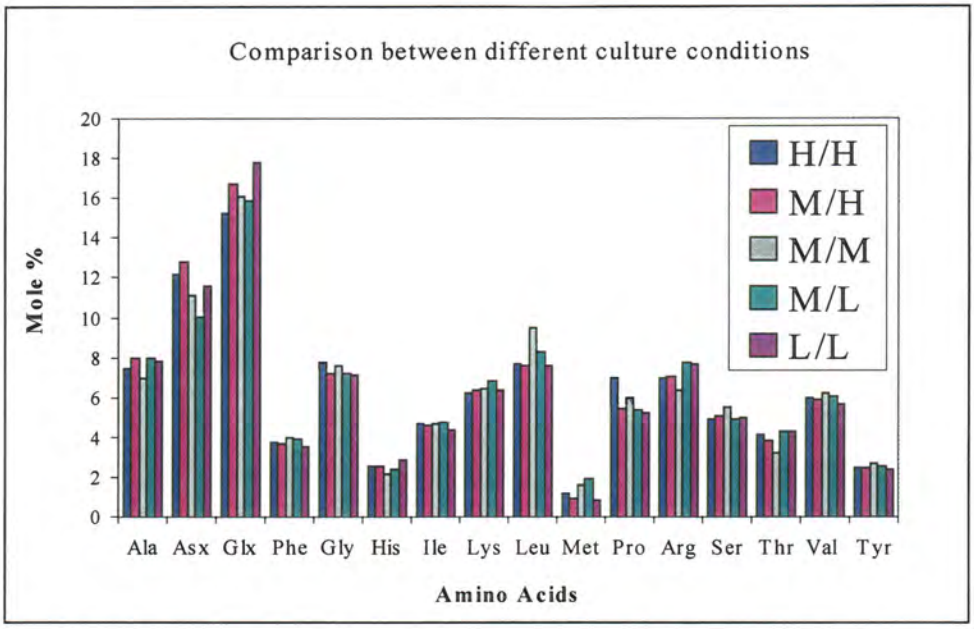


Figure 4.2.1 Comparison of proteinogenic amino acid profiles for the five temperature treatments (* indicates significant difference with respect to M/M temperature treatment)

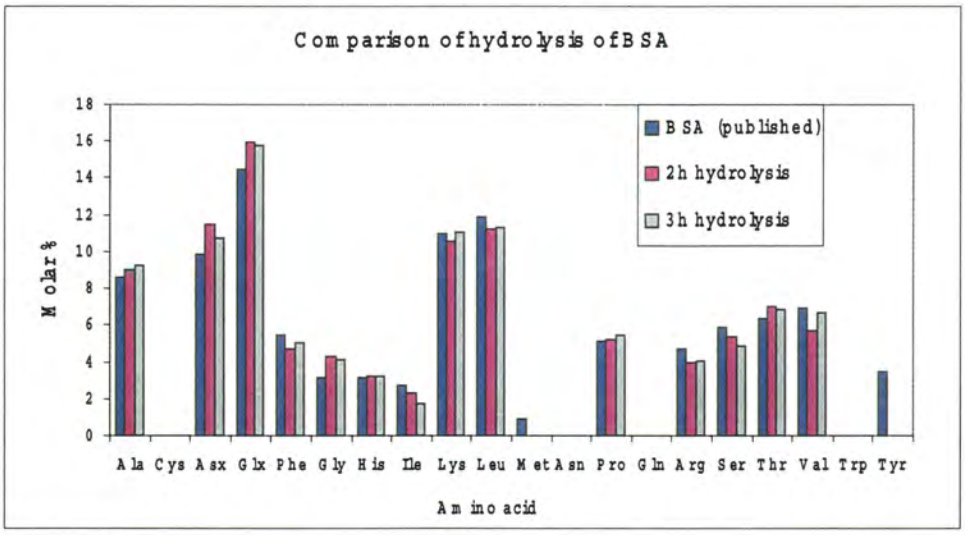


Figure 4.2.2 Comparison of amino acid profiles in Bovine Serum Albumin (BSA) obtained after 2 and 3 hours hydrolysis treatments and published data [47]

analyzed are proteinogenic and the total protein content was not exactly determined. However, the proteinogenic amino acids reflect the total protein content and can be used as important measurements for the flux analysis. As such, the flux profile cannot be directly obtained from the amino acid profile and needs to be supplemented with NMR analysis.

4.3 Metabolic Flux Analysis

4.3.1 Reaction Network for the metabolism of sucrose in soybean embryos

The reaction network for the central carbon metabolism in soybean has been constructed as shown in Figure 4.3.1. The network consists of the glycolysis pathway (GLY), tricarboxylic acid (TCA) cycle, pentose phosphate pathway (PPP), anaplerotic reactions (ANAX), glyoxylate (GOX) shunt and γ-aminobutyrate (GABA) shunt respectively. The reversibilities in the pentose phosphate pathway and the G6P ↔ F6P reaction were also included. The anaplerotic reactions included were the phosphoenolpyruvate carboxylase (PEP ↔ OAA) and malic enzyme (Mal ↔ Pyr) reactions. The PEP carboxylase reaction mainly takes place in the forward direction and is irreversible [48]. The malic enzyme reaction is in the backward direction and irreversible [49]. The reversibilities of the anaplerotic reactions were set to zero accordingly. Since succinate is a symmetrical molecule, there is an inversion of labeling in the succinate to malate reaction in the TCA cycle. This inversion of labeling was incorporated as two fluxes in the system, one with the original carbon skeleton rearrangement and the other with the inverted rearrangement.

In the reaction network representing soybean metabolism, there are 58 reactions and 22 metabolites. Thus, there are F=36 degrees of freedom prior to experimentation. There are 9 external measurements (A-CoA to lipids, P5P, E4P, 3-PG, PEP, Pyr, A-CoA, OAA and Glu to corresponding amino acids respectively). The limits for the substrate consumption (sucrose and glutamine) and effluxes (pyruvate, succinate, P5P efflux, F6P efflux and

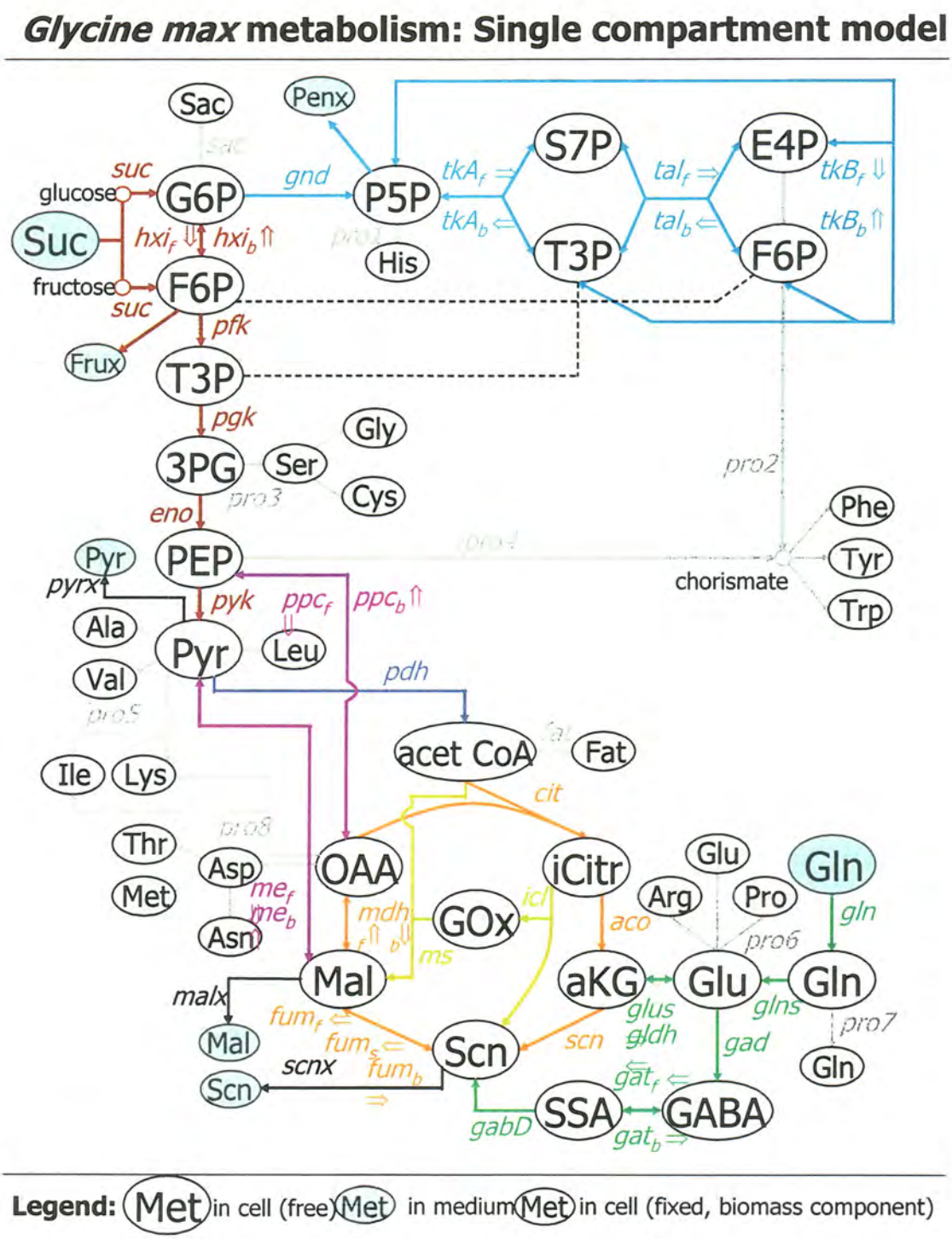


Figure 4.3.1 Single compartment model for the metabolism of sucrose in soybean cotyledons (Appendix C for complete listing of the reactions in the network)

saccharides efflux) were estimated. From the NMR analysis, the intensity fractions for each carbon atom of every amino acid were obtained. The NMR data provided the additional constraints to solve the reaction network. Thus, the system was overdetermined with redundant information, which helped in efficient flux analysis. Appendix C gives the list of all the reactions in the network.

Figure 4.3.1 shows that the amino acids are distributed across the reaction network. Thus, amino acids provide well distributed NMR information. Hence, amino acid measurements were used to quantify the intracellular fluxes. Once the biochemistry and the carbon skeleton rearrangements were identified, the information was used to calculate the external fluxes. The external fluxes and other input parameters were standardized against 20 times glycine production per day as explained in the following section.

4.3.2 Calculation of external fluxes

The external fluxes were determined from the amino acid HPLC analysis, coupled with the precursor-amino acid stoichiometry. The external fluxes, which were calculated and input to the program, are listed in Table 4.3.1. To elucidate, consider Flux 36. OAA is a precursor for Asx, Ile, Lys, Met and Thr. The biochemistry of the pathways indicate that 1 mole of Asx, Ile, Met and Thr each requires 1 mole of OAA. 1 mole of Lys, however, requires 2 moles of OAA. Similar analysis could be applied to the other fluxes as well. Table 4.3.2 shows the relationship of the stoichiometry between metabolites and corresponding amino acids. The amount of protein present could be calculated from the biomass dry weight and the percentage of protein in the biomass for each temperature treatment. The amino acid analysis provides the mol % of the amino acids. Hence, the corresponding moles of each amino acid synthesized to make protein for each temperature treatment could be calculated. The moles of each amino acid were calculated and summed to give the total moles per day of

the precursor metabolite in the system. In the case of Flux 30, it was assumed the nucleotides were 5 % of the total biomass.

Table 4.3.1 External fluxes to be input to the program

Flux #	Precursor metabolite → Amino acids/ Lipids
29	A-CoA → Lipids
30	P5P → His + Nucleotides
31	E4P → Phe + Tyr
32	3-PG → Gly + Ser
33	PEP → Phe + Tyr
34	Pyr → Ala + Val + Ile + Lys + Leu
35	A-CoA → Leu
36	OAA → Asx + Ile + Lys + Met + Thr
37	Glu → Glx + Pro + Arg

The calculation of moles of lipids in the system was carried out as follows. The lipid percentage was known for each temperature treatment. Palmitic acid was assumed to be the major lipid present and the molecular weight of palmitic acid was considered for the calculations ($C_{51}H_{98}O_6$, Molecular weight = 807.34).

$$24 \text{ A-CoA} + 1 \text{ T3P} \Rightarrow 1 \text{ Lipid}$$

Thus, corresponding amount of A-CoA and T3P can be calculated.

The calculated values of precursor metabolites were then standardized with respect to 20 times glycine production per day (Table 4.3.3). The standardization was done to ensure the values of the input fluxes were in the range of 0-2.0. This constraint helped the program to converge by eliminating the occurrence of floating point overflow and underflow errors.

Table 4.3.2 Molar ratios of metabolites to corresponding amino acids

Precursor metabolite ⇒ AA ⇓	3-PGA	Pyr	PEP	E4P	A-CoA	OAA	2-OxoG	R5P
Ala	0	1	0	0	0	0	0	0
Asx	0	0	0	0	0	1	0	0
Glx	0	0	0	0	0	0	1	0
Phe	0	0	2	1	0	0	0	0
Gly	1	0	0	0	0	0	0	0
His	0	0	0	0	0	0	0	1
Ile	0	1	0	0	0	1	0	0
Lys(50%)	0	1	0	0	0	1	0	0
Lys(50%)	0	1	0	0	0	1	0	0
Leu	0	2	0	0	1	0	0	0
Met	0	0	0	0	0	1	0	0
Pro	0	0	0	0	0	0	1	0
Arg	0	0	0	0	0	0	1	0
Ser	1	0	0	0	0	0	0	0
Thr	0	0	0	0	0	1	0	0
Val	0	2	0	0	0	0	0	0
Tyr	0	0	2	1	0	0	0	0

Table 4.3.3 sums up the external fluxes for all the 5 temperature treatments. There was no general trend observed among the temperature treatments. However, in the case of Flux 35 (acetyl-CoA → leucine), the flux value for the M/M temperature treatment was the highest, which is a direct representation of the amino acid analysis where the percentage of leucine was highest for the M/M temperature treatment. Flux 29 (acetyl-CoA → lipids) and 34 (pyruvate → amino acids) followed the trend of the protein and lipid quantities in the five temperature treatments and were highest for the M/H treatment and decreased with decreasing temperatures. Thus, the numerical values of the external fluxes were a combined representation of the biomass dry weight, corresponding percentages of proteins and lipids, calculated sucrose consumption rates and the percentages of the amino acids in the system.

Table 4.3.3 External flux values input to the program for the 5 different temperature treatments (values are normalized relative to 20 times glycine production per day)

Flux # *	H/H	M/H	M/M	M/L	L/L
Glycine (moles/day)	1.48E-6	1.48E-6	1.26E-6	5.06E-7	4.79E-7
29	1.5876	1.6999	1.2632	1.1909	1.1339
30	0.1444	0.1417	0.1230	0.1342	0.1511
31	0.0404	0.0428	0.0438	0.0441	0.0421
32	0.1480	0.1561	0.1392	0.1337	0.1324
33	0.0808	0.0856	0.0876	0.0882	0.0841
34	0.2953	0.3203	0.3269	0.3326	0.3184
35	0.0498	0.0530	0.0627	0.0570	0.0534
36	0.1833	0.1986	0.1792	0.1920	0.1933
37	0.1888	0.2035	0.1876	0.2002	0.2157

* Table 4.3.2 for description of pathways for each Flux #

4.3.3 Calculation of the sucrose and glutamine consumption rates

The culture media at the end of 6 days incubation contained unused sucrose and glutamine as well as glucose, fructose, and other secreted metabolites. Since the sucrose intake and glutamine intake were not measured directly, an alternative method was used to calculate the intake rate for sucrose using data from the literature. The conversion factors for the glucose consumed per kg of product for the major biomass constituents at room temperature were obtained from literature [50]. These factors were considered applicable to the M/M temperature treatment and are tabulated in Table 4.3.4.

Table 4.3.4 Conversion factors for estimating glucose consumed per kg of biomass constituent produced [50]

Biomass components	Glucose consumed (kg glucose/kg product)
Carbohydrates	1.211
Proteins	1.793
Lipids	3.03

Because the amount of protein and % of each amino acid in the protein were known for the M/M temperature treatment, the total carbon (C) content in the protein could be calculated. Similarly, C content in starch and T3P and Acetyl-CoA of the lipids could also be calculated. The amount of cell wall was assumed to be approximately 30 % of the biomass. The combination of starch and cell wall gives the total percentage of carbohydrates in the biomass respectively. Since the exact composition of the cell wall was not known, it was assumed to contain equal percentages of cellulose and hemicellulose. The monomer constituting the cellulose and hemicellulose were considered to be $C_6H_{12}O_6$ and $C_6H_{10}O_5$, respectively. The values for C content in cell wall calculated considering equal percentages

of cellulose and hemicellulose corresponded very well with the C content calculated using cell wall composition data for soybean from literature [51]. Thus, a combination of C content in the major biomass constituents, i.e. protein, starch and carbohydrates was estimated. Using the conversion factors in Table 4.3.4, the amounts of glucose consumed and hence the sucrose consumed was estimated for the M/M temperature treatment. The difference between the C content in sucrose consumed and C content in the biomass gave the C evolved during respiration. The method of calculating sucrose consumption was extrapolated to other temperature treatments and was carried out using two different methods. First, we assumed constant yield of biomass (defined as ratio of the biomass dry weight/ total amount of sucrose consumed) for all the temperature treatments. Because the biomass growth rate per day was known, scaling factors could be obtained for each temperature treatment. The amount of CO_2 evolved was then calculated for the remaining temperature treatments. The amounts of protein, starch and carbohydrates were known for these temperature treatments and the C content in the biomass could be estimated. A sum of the C content in the biomass and the C evolved during respiration estimated the total C content in sucrose consumed for the remaining temperature treatments.

The second method involved considering the temperature coefficient of respiration (Q_{10}) equals 2 for the soybean cotyledons. A Q_{10} of 2 was equivalent to considering that the respiration rate increased or decreased two-fold for a 10°C rise or fall in temperature. The following equation was used to estimate the CO_2 evolution rates for the H/H, M/H, M/L and L/L temperature treatments respectively with respect to the M/M temperature treatment.

$$\ln Q_{10} = (10/T_2\text{-}T_1) \ln (K_2/K_1)$$

where, T_2 is the higher temperature, K_2 is the reaction rate at T_2; T_1 is the lower temperature, K_1 is the reaction rate at T_1 and K is the respiration rates for the respective temperature. For soybean cotyledons, we used the growth rate of the biomass per day at the respective temperatures as a representation of the respiration rates.

Table 4.3.5 summarizes the sucrose consumption rate calculated using both methods. The values reported in Table 4.3.5 were standardized with respect to 20 times glycine production per day. The values estimated using both methods were in the same range. As it was not explicitly known how much cell wall growth takes place during the 6 days in the incubator chamber, sucrose values were calculated using both methods assuming complete cell wall growth or no cell wall growth. Such an assumption helped us to set upper and lower limits for the sucrose consumption input data to the program. Out of the 2 methods used to calculate the sucrose consumption rates, the lower limit and upper limit were set according to the lower of the 2 values and higher of the 2 values. The calculations for the sucrose consumption rate using constant yield and $Q_{10} = 2$ methods are shown in Appendix D. The values of the sucrose intake decrease with decreasing temperature treatments.

The glutamine consumption was calculated using the proteinogenic amino acid measurements. The glutamine fed into the system is primarily used as a nitrogen source for the biomass growth. Consider an amino acid, alanine, its molecular formula is given by $C_3H_7NO_2$. The molar percentage of nitrogen in this amino acid could be calculated from the molecular and atomic weights of alanine (89) and nitrogen (14) as (14/89) *100, which is equal to 15.73 %. Similarly, the percentage of nitrogen in other amino acids could be calculated and the total amount of nitrogen in amino acids was estimated. The nitrogen percentages could be converted to the total amount of glutamine consumed in the system. The insoluble protein percentage ranged from 2-5 % of the biomass dry weight. To encompass the entire range of glutamine intake, the limits for the intake were set as shown below in Table 4.3.6. The absolute values of glutamine intake decreased with decreasing temperature, which was due to the lower amount of protein in the lower temperature treatments. Thus, it was concluded that the capacity of synthesis controls the substrate uptake by the soybean cotyledons. The analysis of the NMR data is discussed in the next section.

Table 4.3.5 Standardized sucrose consumption rates input to the program for five temperature treatments

Calculation Method	Limits	H/H	M/H	M/M *	M/L	L/L
Constant yield	Lower	**0.94**	**0.97**	**0.84**	**0.87**	**0.88**
	Upper	1.27	1.29	1.13	1.17	1.22
$Q_{10} = 2$	Lower	0.96	0.99	0.84	0.90	0.91
	Upper	**1.37**	**1.41**	**1.13**	**1.33**	**1.37**

* Limits calculated for this treatment using both methods are the same for both methods since M/M is the standard treatment.
Bold faced values are the final lower and upper limits set for the program

Table 4.3.6 Standardized glutamine consumption rates input to the program for five temperature treatments

Glutamine intake limits	H/H	M/H	M/M	M/L	L/L
Calculated (moles/day)	**1.39E-5**	**1.52E-5**	**1.19E-5**	**5.15E-6**	**5.04E-6**
Calculated *	**0.59**	**0.55**	**0.47**	**0.66**	**0.53**
Lower *	0.52	0.50	0.40	0.57	0.47
Upper *	0.70	0.60	0.55	0.70	0.60

* Values normalized with 20 times glycine production per day

4.3.4 NMR data analyses

The signal to noise ratio (S/N) in an NMR experiment depends directly on the concentration of the nuclei (C) and the number of scans (N_s).

$$S/N \propto C * (N_s)^{0.5}$$

The accumulation time (T_a) for a free induction decay (FID) depends on N_s and T_P, where T_p is given by the following equation [52].

$$T_p = t_p + T_{acq} + T_{rd}$$

where, t_p is the length of radiofrequency pulse, T_{acq} is the acquisition time and T_{rd} is the relaxation delay. And,

$$T_a = T_P * N_s$$

Thus, to double the signal to noise ratio in the case of low concentrations, the number of scans needs to be increased four times. Since T_a depends directly on N_s, it will also increase proportionately. Hence, a balance is essential between T_a, N_s and S/N as the cost required to run an NMR experiment is very expensive.

The protein quantities shown in Table 4.2.1 indicate that the concentration was the highest in the case of M/H temperature treatment. The H/H and M/M also had high amounts of protein whereas the protein concentration was low for the L/L and M/L temperature treatments. In the case of M/H and M/M temperature treatments, 2 out of 3 samples were pooled to increase the protein concentration for NMR analysis. In the case of H/H temperature treatment, all the 3 samples were pooled to reduce NMR run time. In the case of M/L and L/L temperature treatments, a single sample was taken and the number of scans in the experiment was increased to compensate for the concentration, which increased the NMR time.

The 2D NMR analyses were carried out differently to achieve a balance between NMR time and the concentration of the protein in the NMR tube. Only 10 % labeled sucrose was fed to the soybean cotyledons. If this 10 % labeling randomly distributes through the

network, the probability that two adjacent atoms are labeled, is about 1 % and the probability that 2 adjacent atoms originating from the same metabolite being labeled is 10% [6]. The minimum concentration for a 2D NMR analysis is 1mM, from the specifications of a 500 MHz spectrometer and the amino acid with the lowest concentration was methionine (2 mole %). Hence, for approximately 20-22 hour NMR experiment, taking all the above mentioned factors into consideration the minimum amount of protein required in our sample for the 500 MHz spectrometer was 7 mg.

The two analyses carried out on the samples of five temperature treatments were 2D-HSQC (Heteronuclear Single Quantum Correlation spectroscopy) and 2D-TOCSY (TOtal Correlation SpectroscopY). The spectrum of the 2D-HSQC for the M/M sample is shown in Figure 4.3.2. From the known chemical shift and J coupling data, the peaks could be identified on the spectrum. The NMR software (NMRview) enabled a 1D slice of the peak and the area under the peaks could be quantified. For example, the magnified peak of the spectrum as shown in Figure 4.3.2 was proline-δ. From the known biochemistry of the pathway, it was deduced that in the case of pro-δ, the multiplets observed would be a singlet and a doublet. The 1D slice of pro-δ shows the singlet and doublet peaks. The intensities obtained from the area under the peak were used to report the relative fractions of the singlet and the doublet with respect to the total intensity for pro-δ peak. The noise was also calculated in the different regions of the spectrum. Figure 4.3.3 (a) shows the alpha peak of aspartate/ asparagines. In the case of the alpha peak, we should observe a singlet, a smaller doublet, a bigger doublet and a doublet of a doublet (Figure 2.4.1). Thus, in all 9 peaks should be observed. However, 7 out of the 9 peaks are visually clearly observable. It was inferred that since the chemical shifts are in close range of each other some peaks might overlap. As a result, some of the observed peaks were actually a combination of more than one peak. This overlap made deconvolution of the peaks necessary. A computer program for deconvolution was developed. Figure 4.3.3 (b) shows that the program works effectively in

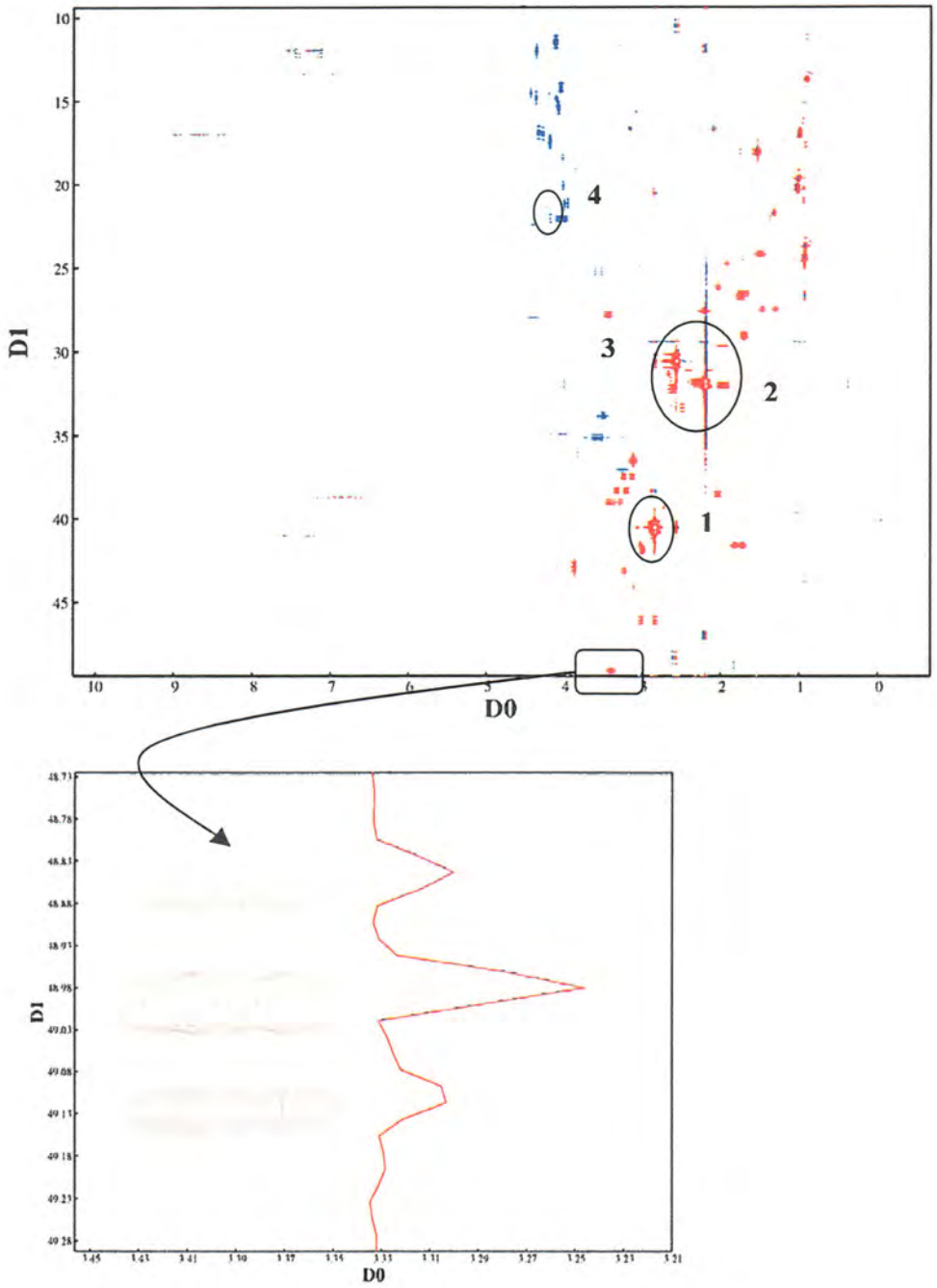
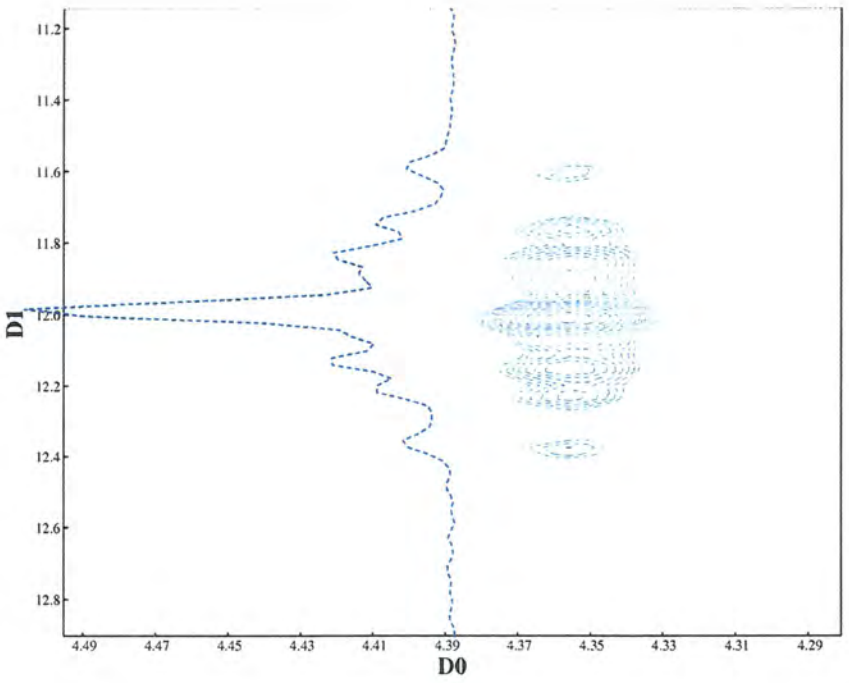


Figure 4.3.2 2D HSQC spectrum for amino acids from the H/H temperature treatment where D1 is the ^{13}C chemical shift and D0 is the ^{1}H chemical shift. The enlarged section represents the proline-δ peak and its 1D slice from the 2D HSQC spectrum. (1, 2, 3 and 4 are peaks of hydrolyzed sugar molecules explained in section 4.5.4)

(a)



(b)

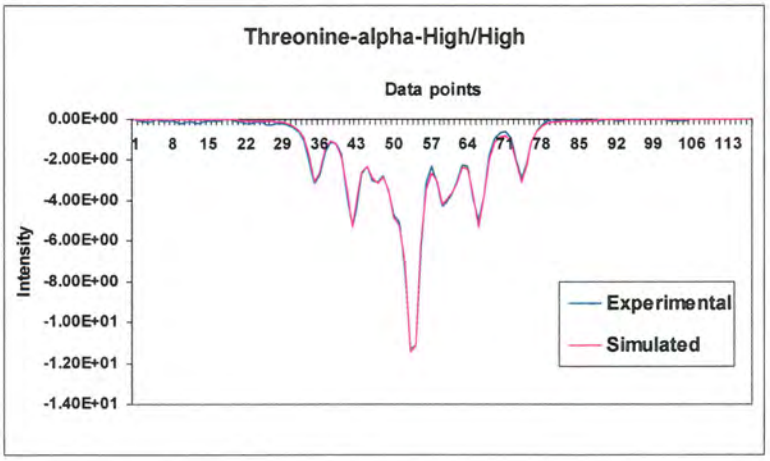


Figure 4.3.3 (a) 1D slice of aspartate-α from the 2D HSQC spectrum from H/H temperature treatment, where D1 is the ^{13}C chemical shift and D0 is the ^{1}H chemical shift (b) Peak fitting of threonine-α using the deconvolution program

fitting the experimental and simulated data.

The TOCSY spectrum shown in Figure 4.3.4 was used to calculate the carbon enrichments. The magnified peak shown in the Figure 4.3.4 belongs to Phenylalanine-ε. The TOCSY analysis detects the protons attached to ^{12}C and ^{13}C. The central peak for Phe-ε was stronger compared to the satellite peaks on either side and signified the protons attached to the ^{12}C. The satellite peaks were from the protons attached to the ^{13}C and thus denoted the enrichment of the carbon atom under consideration. If each satellite peak is denoted by b and the central peak is denoted by a, the enrichment can be calculated as

$$\text{Enrichment} = 2b/(a+2b)$$

The combined data of the HSQC and TOCSY experiments was reported for each carbon atom for the proteinogenic amino acids. The NMR data provide additional constraints for the system. The external fluxes calculated combined with the NMR data form the input data to the computer program, which calculated the intracellular fluxes.

Table E.1 in Appendix E gives a comprehensive listing of the entire relative intensity fraction for the carbon atoms of the amino acids for all the different temperature treatments. The comparisons between the temperature treatments are discussed in detail in section 4.5.2.

4.3.5 Additional input data assigned to the program

Apart from specifying the sucrose and glutamine intake limits, limits for additional parameters were assigned to the program. The other parameters include the pyruvate efflux, succinate efflux, penx (efflux from P5P) and frux (efflux from F6P) and saccharides (sac, efflux from G6P) limits. Since the pyruvate and succinate efflux were not measured in our experiment, the upper limits were set taking into consideration the possibility of an efflux. The upper limits for the penx and the frux were set using values from the literature [51]. Table 4.3.7 summarizes the limits for these additional parameters for the five treatments. The mathematical model and the simulations output are discussed in the following section.

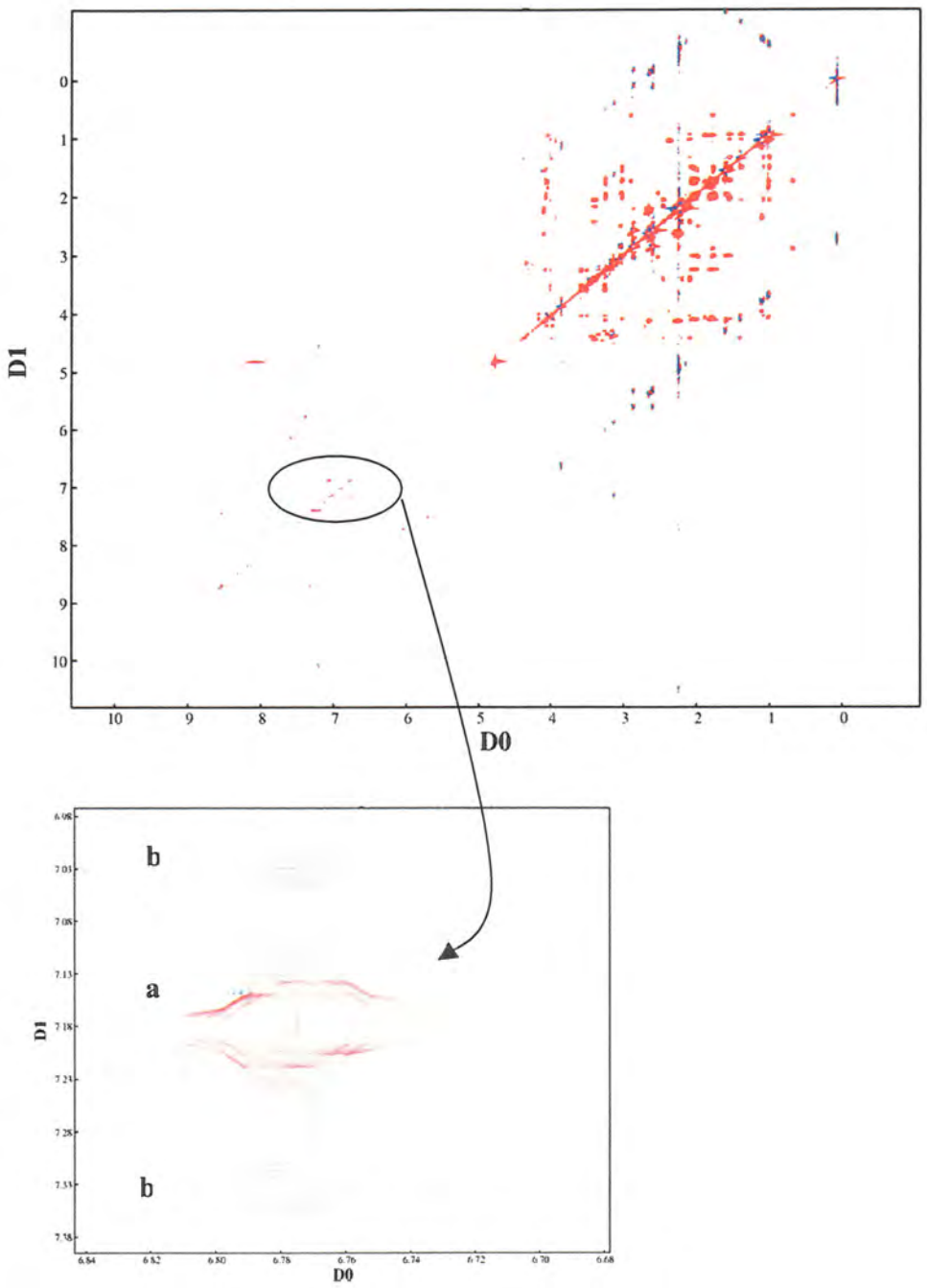


Figure 4.3.4 2D TOCSY spectrum of the H/H temperature treatment where D1 and D0 are the chemical shifts of ^{1}H. The enlarged portion represents the peak of Phe-ε and its 1D slice from the 2D TOCSY spectrum. (a and b indicate the central and satellite peaks respectively)

Table 4.3.7 Summary of additional input data to the program for the five temperature treatments

Parameters	H/H	M/H	M/M	M/L	L/L
Pyruvate efflux	0-0.5	0-0.5	0-0.5	0-0.5	0-0.5
Succinate efflux	0-0.8	0-0.8	0-0.8	0-0.8	0-0.8
Sac	0.10-0.35	0.10-0.35	0.10-0.35	0.14-0.35	0.11-0.35
Penx	0-0.35	0-0.35	0-0.35	0-0.35	0-0.35
Frux	0-0.02	0-0.02	0-0.02	0-0.02	0-0.02

Sac represents saccharides, penx represents efflux from pentose-5-phosphate and frux represents efflux from fructose-6-phosphate

4.4 Mathematical model for the reaction network

4.4.1 Algorithm for evaluating intracellular fluxes in soybean embryos

To solve the problem of converting the NMR data to flux data, an initial set of guessed fluxes was used to generate a simulated isotopomer distribution. The isotopomer distribution was compared to the experimental distribution from the NMR data and the process can be reiterated until the error between the simulated and experimental distribution was minimized. This approach uses two innovative techniques; a singular value decomposition method to obtain a feasible set of fluxes and a Boolean function mapping method to enumerate the isotopomers for the system. The Boolean function mapping method requires minimal input from the user. The mathematical model also incorporated all the known reversible reactions. This was important as the reversibility could affect the flux values to a large extent especially in the case of complex plant systems. The model uses a simulated annealing method for error minimization and a Monte Carlo method for refining guesses. Once the global optimum for the error between the simulated and experimental NMR intensities was found, a quadratic search (Powell method) was used to converge closer

to the real optimum. The statistical analysis of the intracellular fluxes obtained was carried out using a constant χ^2 method, which gave 90 % confidence intervals for the fluxes [9].

The zeroth run of the simulation was carried out using the actual values of the NMR intensity fractions input to the program. Five hundred simulations were then carried out by perturbing each data point randomly using a normal distribution within the limits of the NMR error. For example, the Ile γ^1 peak had a singlet fraction of 0.2519 and a doublet fraction of 0.7481 with an NMR experimental error of 0.045. Thus, the values of the singlet and double fractions were perturbed in the region of ±0.045 errors randomly using a normal distribution. A probability distribution function was plotted for important fluxes and flux ratios in the metabolic pathway. The above mentioned approach was used for the five temperature treatments.

Consider a reversible reaction, A + B ↔ C + D. If the flux in the forward reaction is represented by v + r and the reverse flux is r, then E = r/ (v+r) is the extent of reversibility. If E is zero, it indicates that the reaction is irreversible and if E is 1, it indicates that there is no net flux in the forward direction. The limits of E were set as $E_{min} = 0$ and $E_{max} = 0.999$ for all the reversible reactions in the network. For an E_{max} value of 0.9999, r (the reverse flux) can take very large values. A very large value of r, however, is impractical as the enzymes catalyzing the reversible reactions will not be able to handle such high fluxes. The optimum to which the program converged was the same for $E_{max} = 0.999$ and Emax = 0.9999. However, for an E_{max} value of 0.99, the optimum was different. Thus, it was concluded that the E_{max} value of 0.999 is a good estimate for flux analysis. Figure 4.4.1 shows the probability distribution for the PPP to glycolysis flux ratio for the M/M temperature treatment. It can be seen from Figure 4.4.1 that the range of different flux values taken or the flux ratio (that the program converges to) is tighter for simulations run with E_{max} value of 0.999 as compared to the E_{max} value of 0.9999. A comprehensive analysis of the results of the simulations is discussed in the following section.

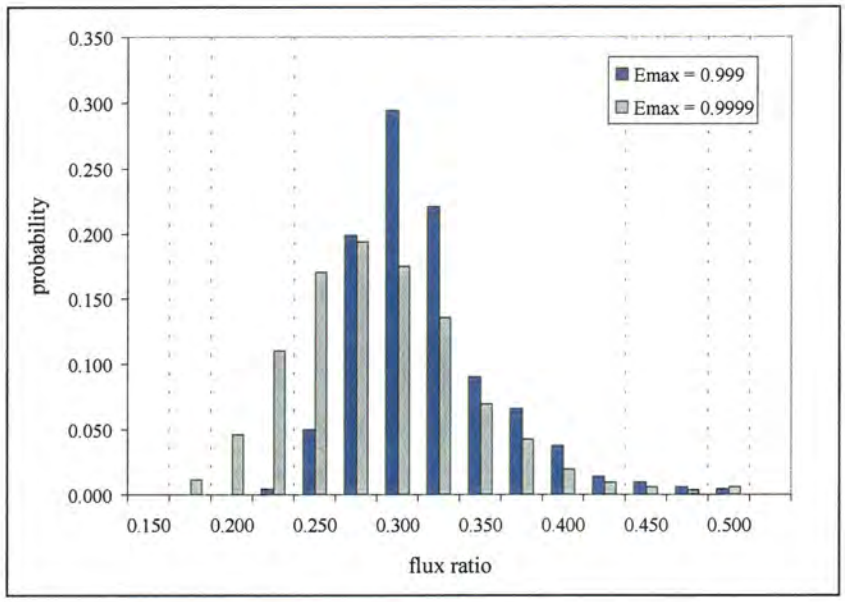


Figure 4.4.1 Probability distribution for Pentose Phosphate Pathway/ Glycolysis flux ratio evaluated at two levels of E_{max} for M/M temperature treatment

4.4.2 Program output for the M/M temperature treatment

The external fluxes and NMR data were used as inputs for 501 simulations, which were carried out for the metabolic flux analysis in the developing soybean cotyledon for the M/M temperature treatment. The zeroth simulation took around 5.4 minutes and the remaining 500 simulations took an average of 1.9 minutes each. Thus, for a particular temperature treatment, these 501 simulations took a total time of 16 hours. Figure 4.4.2 represents a metabolic map of the central carbon metabolism in developing soybean cotyledons. The flux values shown in the map are standardized with respect to 20 times glycine production per day.

The metabolic flux map can be used as an important tool for making important observations about pathway functioning. A probability density function was plotted for the PPP to glycolysis and PPP to sucrose intake flux ratios for the 500 simulations. The flux of

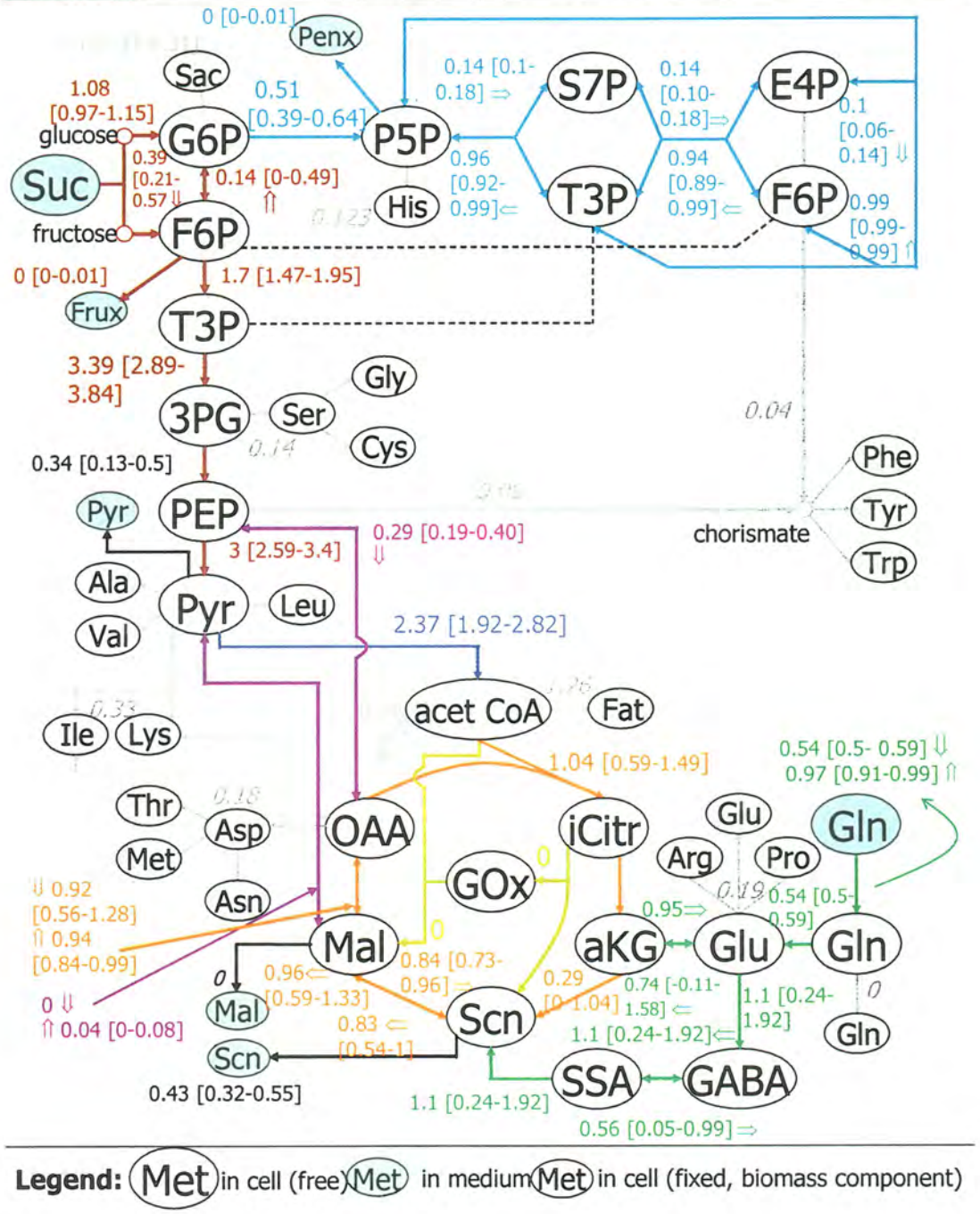


Figure 4.4.2 Metabolic flux map for developing soybean cotyledon for the M/M temperature treatment using a single compartment model

the carbon entering the PPP pathway from G6P represents one carbon lost as CO_2 and the remaining 5 carbons as P5P. F6P metabolizes to two T3P molecules, which doubles the flux from F6P to T3P. Taking these metabolic steps into consideration, the PPP to glycolysis flux ratio was calculated as PPP flux/ (0.5 * T3P flux).By the same analogy, the PPP to sucrose intake ratio was calculated as PPP flux/ (2 * Sucrose intake). The probability density functions for these flux ratios are shown in Figure 4.4.3. The mean value of the PPP/Glycolysis flux was 0.30 with a 90% confidence interval of [0.23, 0.38]. The mean value indicates that the PPP flux is around 30 % of the glycolysis flux. For the PPP/ Total sucrose intake ratio, the mean value was 0.24 with a 90% confidence interval of [0.18, 0.29]. This value indicates that 24 % of the sucrose input flux flows into the pentose phosphate pathway. The extent of reversibilities for the transketolase and transaldolase enzymes catalyzing the PPP reactions were 0.97 [0.92, 0.999], 0.94 [0.90, 0.999] and 0.999 [0.999, 0.999] respectively. Thus, the reversible reactions in the PP pathway reactions were found to have no net flux. The extent of reversibility for the G6P $\leftrightarrow$ F6P reaction had a mean of 0.14 with a 90 % confidence interval of [0-0.5]. This low value, though not very well determined, indicates that the G6P $\rightarrow$ F6P reaction was not very reversible in the case of the M/M temperature treatment. The anaplerotic reaction PEP $\leftrightarrow$ OAA catalyzed by the PEP carboxylase enzyme had a relative flux (with respect to glycine) of 0.31 [0.2, 0.4]. The flux through the anaplerotic reaction Pyr $\leftrightarrow$ Mal catalyzed by the malic enzyme was a negligible flux. Figure 4.4.4 shows the probability distributions for the flux ratios of PEP carboxylase to pyruvate dehydrogenase (ppcf/pdh), malic enzyme to pyruvate dehydrogenase (meb/pdh) and the net anaplerotic reactions to pyruvate dehydrogenase (ana/pdh) reactions. The ratio of the pyruvate dehydrogenase reaction to the PEP carboxylase reaction was 10 indicating that the most of the flux channels into lipids and the TCA cycle. Thus, down regulating the PEP carboxylase gene in soybean would not help in increasing the carbon partitioning, say in lipids as the flux through PEP carboxylase is very small as compared to flux through

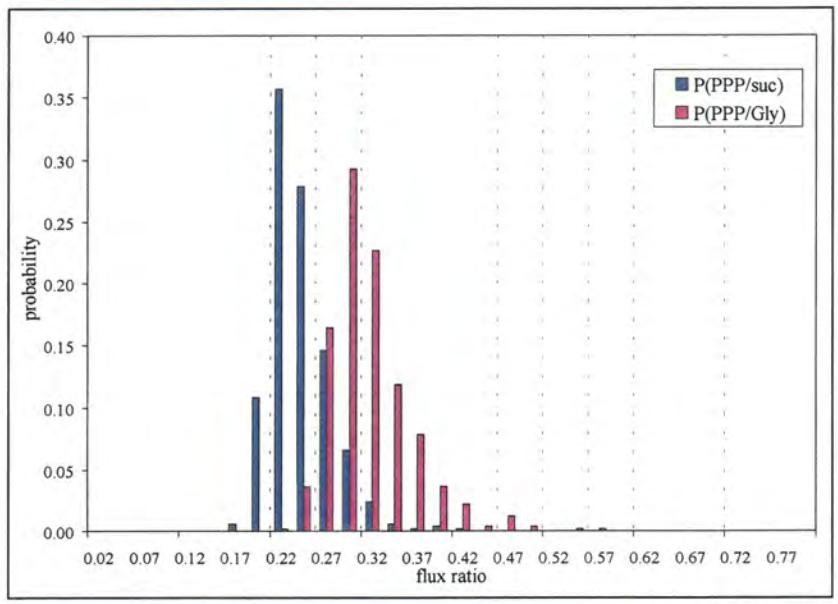


Figure 4.4.3 Probability distribution for flux ratios of PPP/Glycolysis and PPP/Sucrose intake for the M/M temperature treatment

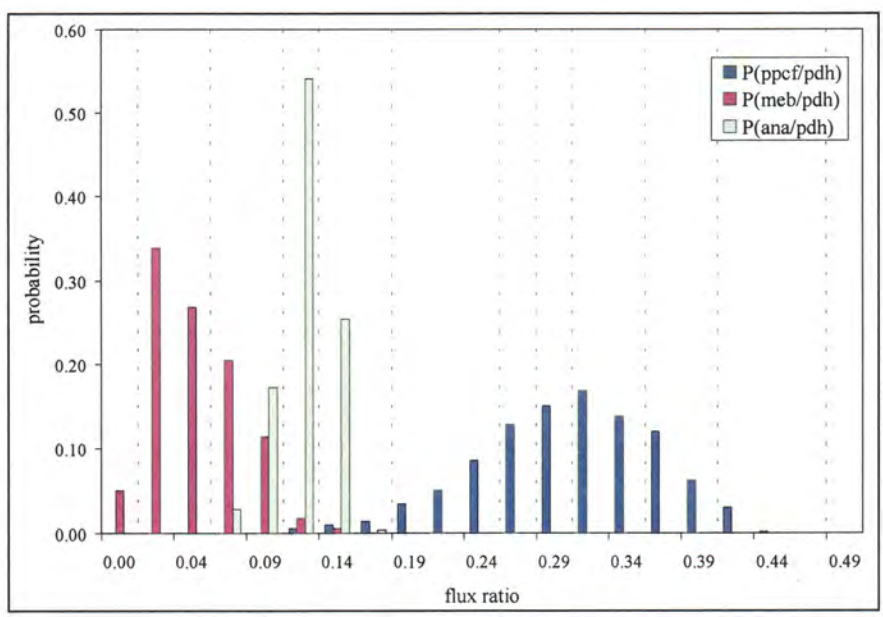


Figure 4.4.4 Probability distribution for flux ratios for M/M temperature treatment (ppcf indicates PEP-carboxylase enzyme catalyzing the forward reaction of PEP to OAA, meb indicates the malic enzyme catalyzing the backward reaction of Mal to Pyr and pdh indicates pyruvate dehydrogenase enzyme catalyzing the reaction of Pyr to A-CoA)

pyruvate dehydrogenase.

The glyoxylate shunt was found to be inactive. The limits of succinate and pyruvate effluxes were set to 0-0.8 and 0-0.5, respectively, for the M/M temperature treatment. Figure 4.4.5 shows the probability density function for pyruvate and succinate efflux. The mean value of the succinate efflux was 0.44 with a 90 % confidence interval of [0.32, 0.55]. However, the mean pyruvate efflux was 0.34 but the standard deviation from the 500 simulations was very high (0.13) indicating that this flux was not well determined. A trial run was carried out by keeping all the other values of the input parameters constant and changing the pyruvate efflux limit from 0-0.5 to 0-2.0. The standard deviation was very high even in this case and program converged to pyruvate efflux values ranging from 0-2.0. Thus, the pyruvate efflux was not 'identifiable' for the soybean system. The reaction network is represented with a mathematical model supplemented with external measurements. The information about different parts of the pathway is obtained from measurements pertaining to that particular part of the pathway. For example, flux analysis of the reversibilities of the transketolase and transaldolase reactions in the pentose phosphate pathway depend primarily on the labeling information obtained from the PEP family of the amino acids and His. If the external measurements or the independent NMR information are not sufficient to estimate the fluxes or if the error in the measurements is large, then the fluxes become 'structurally unidentifiable'. This problem can be solved by increasing the number of measurements or providing the system with low noise level measurements. However, in some cases, the relationship between the NMR measurements and the fluxes are very complicated. In such cases, the fluxes become 'statistically unidentifiable' and a very low noise level can translate into large confidence intervals for the fluxes. Thus, in the case of a statistically unidentifiable flux, the model cannot estimate the flux irrespective of redundant measurements pertaining to that flux and the pyruvate efflux is one such measurement in our system. Although, succinate efflux was identifiable to a certain extent, direct measurements of pyruvate, succinate and

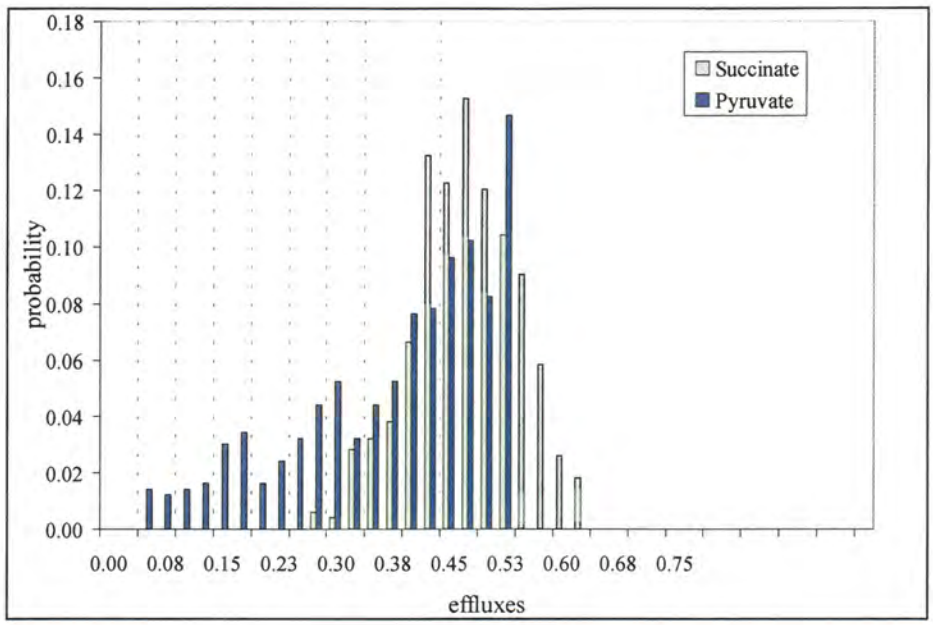


Figure 4.4.5 Probability distribution for efflux rates of succinate and pyuvate from the cultures of soybean cotyledon for M/M temperature treatment

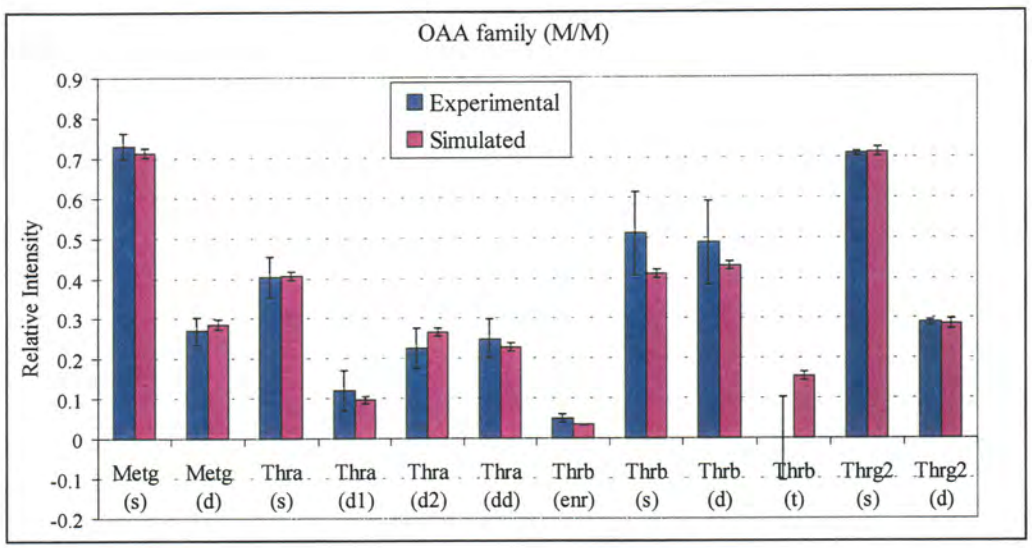
other organic acid effluxes are needed for better estimation of metabolic fluxes for the soybean system.

In the lower part of the reaction network comprising of the TCA cycle, GABA and GOX shunt, it is observed that though there is considerable flux through the GABA shunt, the reactions associated with it are not identified. The reason for non-identifiability could be due to dilution of carbon in that part of the network due to the unlabeled glutamine. This dilution effect results in weak metabolic information from the NMR spectrum for the metabolites associated in lower part of the reaction network. Feeding the system with labeled glutamine in addition to labeled sucrose can possibly solve this issue of non-identifiability. Earlier studies have reported flux evaluation through the GABA shunt in *Aspergillus niger* using labeled glutamine as the feed [53].

4.4.3 Comparison of simulated and experimental NMR intensities for the M/M temperature treatment

The comparison of the simulated and experimental NMR intensity of the carbon atoms of the amino acids was used to test the validity of the single compartmental model constructed for the metabolism of developing soybean cotyledon. For a terminal carbon atom, for example, proline-δ, there are two possibilities. It is attached to proline-γ which is ^{12}C and this gives rise to a singlet peak in the NMR spectrum; or it is attached to a proline-γ which is ^{13}C and this gives rise to a doublet peak in the NMR spectrum. Thus, the singlet and doublet peaks of proline-δ are represented as relative intensities (or fractions) with respect to the total peak intensity. Figure 4.4.6 shows clearly that the simulated intensity compares very well (3-7 % difference) with the experimental intensity for the OAA and pyruvate family of amino acids. Figure 4.4.7 shows the comparison of the intensities for the PEP family of amino acids, Phe and Tyr. In some cases, the difference between the simulated and experimental intensities is more than 10 %. This difference indicates that the model is not valid for the upper part of the pathway comprising the glycolysis and PPP reactions. A possible compartmentation in the upper part of the pathway could be the reason for the difference in the simulated and experimental intensities of PEP family of amino acids. The plant system is known to be complex with most of the reactions in the present model occurring in multiple compartments. The assumption made was that the exchange between the compartments is extremely high and the flux value in the one-compartment model represents a net flux. However, for the upper part of the network comprising of the glycolysis and PPP reactions, this assumption was not valid. Hence, a two-compartment (cytosolic and plastidic) model needs to be incorporated for better estimation of fluxes in soybean cotyledons. The following section 4.5 discusses the effects of temperature in the central carbon metabolism in developing soybean cotyledon.

(a)



(b)

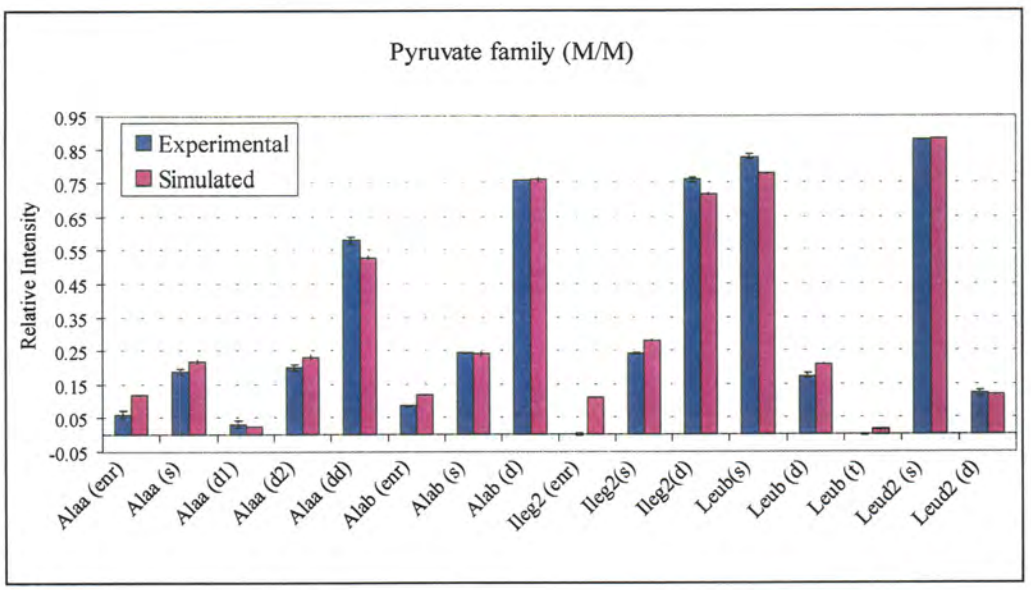


Figure 4.4.6 Comparison of experimental and simulated intensities for (a) OAA family of amino acids (b) Pyruvate family of amino acids for M/M temperature treatment (a, b, d and g indicate ∝, β, δ and γ carbon atom; s, d1, d2 and dd indicates singlet, doublet, second doublet and double doublet respectively)

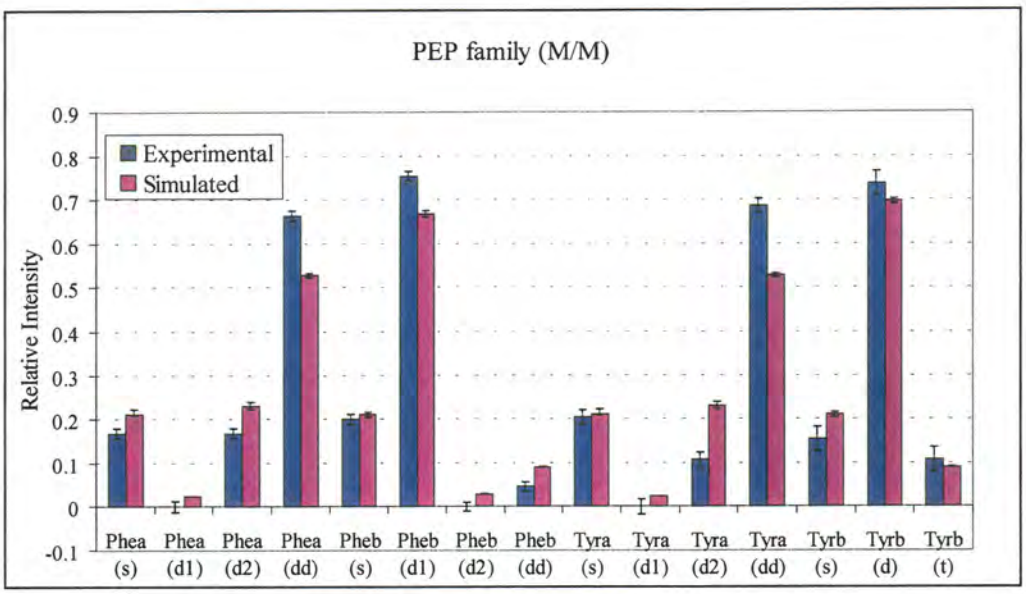


Figure 4.4.7 Comparison of experimental and simulated intensities for PEP family of amino acids for M/M temperature treatment (a and b indicate ∝ and β carbon atom; s, d1, d2, dd and t indicates singlet, doublet, second doublet, double doublet and triplet respectively)

4.5 Temperature effects on soybean metabolism

4.5.1 Comparison of program outputs for the H/H, M/H, M/L, L/L and M/M temperature treatment

Computer simulations were carried out for the remaining four temperature treatments H/H, M/H, M/L and L/L respectively. As in the case of the M/M temperature treatment, the probability distribution was plotted for PPP to glycolysis ratio, PPP to sucrose intake ratio, PEP carboxylase to pyruvate dehydrogenase, malic enzyme to pyruvate dehydrogenase and the net anaplerotic reactions to pyruvate dehydrogenase reactions respectively for the remaining temperature treatments. The student's t-Test was used to compare the flux ratios at different temperature treatments with respect to the M/M temperature treatment using the jmp software (version 5.0.1). The results of this test are shown in Figure 4.5.1. All the flux ratios were significantly different from values in the M/M temperature treatment. However,

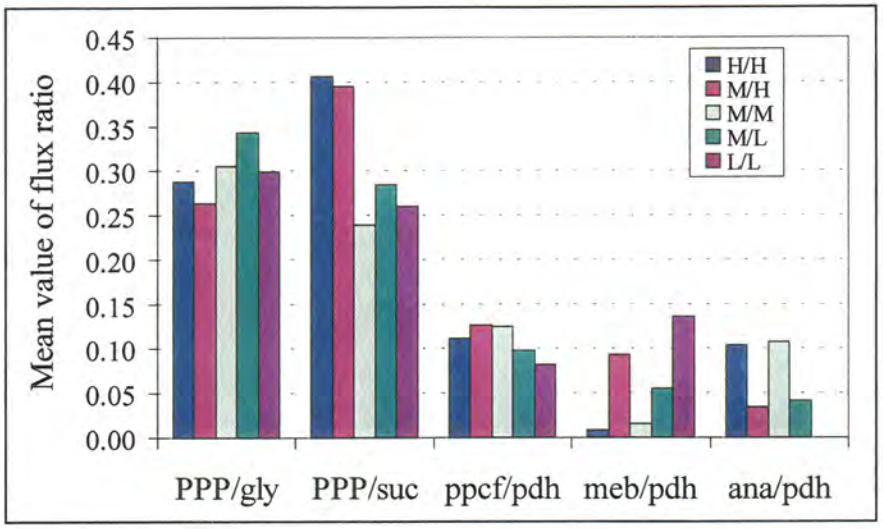


Figure 4.5.1 Comparison of flux ratios of different temperature treatments with respect to the M/M temperature treatment. All these ratios are statistically significantly different except PPP/glycolysis for L/L and M/M and PPP/suc for M/L and M/M temperature treatments respectively. (Flux ratios plotted are pentose phosphate pathway to glycolysis, PPP to sucrose intake, PEP carboxylase reaction to pyruvate dehydrogenase (pdh), malic enzyme reaction to pdh and net anaplerotic reaction to pdh)

the PPP to sucrose intake ratio for the M/L and M/M temperature treatments and the PPP to glycolysis reaction for the L/L and M/M temperature treatments were not statistically significantly different.

Table 4.5.1 lists the absolute values of the metabolic fluxes for all five temperature treatments. The corresponding metabolic flux maps for the remaining four temperature treatments are shown in Appendix F. Table 4.5.2, which shows the absolute values of the fluxes, shows that the sucrose intake decreases with decreasing temperatures and the glutamine intake is highest for the M/H temperature treatment. The protein produced in the M/H was the highest and thus, the flow of the carbon reflected in the pathway agrees with the trend of the amount of protein or oil produced in the system at that particular temperature treatment. The flux ratios are calculated to analyze different nodes in the reaction network

Table 4.5.1 Absolute flux values calculated for the reactions for all five temperature treatments

Reaction	Enzymes	H/H	Confidence Interval (90%)		M/H	Confidence Interval (90%)		M/M	Confidence Interval (90%)		M/L	Confidence Interval (90%)		L/L	Confidence Interval (90%)	
#		Mean	Lower	High	Mean	Lower	High	Mean	Lower	High	Mean	Lower	High	Mean	Lower	High
1	**in1**	1.937	1.745	2.043	1.991	1.820	2.087	1.363	1.229	1.455	0.576	0.479	0.672	0.557	0.449	0.656
2	**in2**	0.737	0.620	0.844	0.789	0.666	0.888	0.686	0.632	0.746	0.262	0.223	0.300	0.275	0.253	0.287
3	**hxif**	0.814	0.590	1.038	0.913	0.712	1.114	0.495	0.272	0.719	0.194	0.006	0.382	0.231	0.041	0.422
4	**hxib**	*0.169*	*0.000*	*0.582*	*0.015*	*0.000*	*0.139*	*0.137*	*0.000*	*0.493*	*0.737*	*0.231*	*0.999*	*0.764*	*0.387*	*0.999*
5	**pgl**	0.915	0.723	1.106	0.867	0.709	1.025	0.651	0.495	0.806	0.252	0.105	0.400	0.216	0.114	0.317
6	**tktAf**	0.253	0.189	0.317	0.239	0.187	0.292	0.183	0.132	0.234	0.032	0.015	0.065	0.024	0.012	0.039
7	**tktAb**	*0.969*	*0.923*	*0.999*	*0.973*	*0.941*	*0.999*	*0.965*	*0.920*	*0.999*	*0.986*	*0.971*	*0.999*	*0.987*	*0.972*	*0.999*
8	**talf**	0.253	0.189	0.317	0.239	0.187	0.292	0.183	0.132	0.234	0.032	0.015	0.065	0.024	0.012	0.039
9	**talb**	*0.941*	*0.893*	*0.989*	*0.963*	*0.931*	*0.994*	*0.943*	*0.896*	*0.990*	*0.981*	*0.960*	*0.999*	*0.985*	*0.968*	*0.999*
10	**tktBf**	0.193	0.130	0.257	0.176	0.124	0.228	0.127	0.076	0.178	0.010	-0.007	0.043	0.003	-0.008	0.019
11	**tktBb**	*0.999*	*0.999*	*0.999*	*0.999*	*0.999*	*0.999*	*0.999*	*0.999*	*0.999*	*0.999*	*0.999*	*0.999*	*0.999*	*0.999*	*0.999*
12	**pfk**	3.196	2.810	3.506	3.319	2.976	3.598	2.166	1.863	2.465	0.806	0.588	1.024	0.809	0.539	1.080
13	**gap**	6.367	5.566	6.928	6.583	5.875	7.107	4.284	3.665	4.851	1.554	1.129	1.978	1.559	1.023	2.095
14	**pyk**	5.704	5.032	6.245	5.782	5.205	6.288	3.795	3.287	4.304	1.391	1.024	1.757	1.403	0.935	1.872
15	**pdh**	4.830	4.072	5.588	5.295	4.600	5.989	2.998	2.425	3.571	1.182	0.806	1.559	1.356	0.792	1.919
16	**cs**	2.372	1.622	3.122	2.695	2.001	3.388	1.318	0.746	1.891	0.548	0.173	0.923	0.782	0.223	1.341
17	**scn**	1.179	0.003	2.665	0.716	0.000	1.998	0.371	0.000	1.327	0.072	0.000	0.318	0.209	0.000	0.675
18	**fum1f**	2.140	1.501	2.778	2.806	2.184	3.428	1.216	0.748	1.685	0.594	0.233	0.955	0.943	0.380	1.507
19	**fum1b**	*0.731*	*0.552*	*0.910*	*0.664*	*0.539*	*0.788*	*0.843*	*0.727*	*0.958*	*0.850*	*0.687*	*0.999*	*0.587*	*0.455*	*0.718*

italics indicate the extent of reversibilities
All fluxes have the unit of moles/day and need to multiplied by a factor of 1E-6

Table 4.5.1 Continued

Reaction #	Enzymes	H/H	Confidence Interval (90%)		M/H	Confidence Interval (90%)		M/M	Confidence Interval (90%)		M/L	Confidence Interval (90%)		L/L	Confidence Interval (90%)	
		Mean	Lower	High	Mean	Lower	High	Mean	Lower	High	Mean	Lower	High	Mean	Lower	High
20	**fum12**	1.396	1.038	1.480	1.378	1.194	1.480	1.061	0.681	1.265	0.467	0.311	0.506	0.423	0.347	0.479
21	**mdhf**	2.117	1.496	2.738	2.317	1.767	2.866	1.168	0.716	1.620	0.528	0.211	0.846	0.762	0.270	1.253
22	**mdhb**	*0.950*	*0.863*	*0.999*	*0.000*	*0.000*	*0.000*	*0.000*	*0.000*	*0.000*	*0.000*	*0.000*	*0.000*	*0.000*	*0.000*	*0.000*
23	**ppcf**	0.543	0.374	0.713	0.675	0.519	0.830	0.378	0.252	0.504	0.118	0.048	0.189	0.115	0.044	0.186
24	**ppcb**	*0.004*	*0.000*	*0.010*	*0.000*	*0.000*	*0.000*	*0.000*	*0.000*	*0.000*	*0.000*	*0.000*	*0.000*	*0.000*	*0.000*	*0.000*
25	**mef**	0.007	0.000	0.015	0.010	0.000	0.015	0.007	0.000	0.013	0.003	0.000	0.005	0.003	0.000	0.005
26	**meb**	*0.027*	*0.000*	*0.082*	*0.332*	*0.245*	*0.420*	*0.039*	*0.000*	*0.087*	*0.132*	*0.000*	*0.330*	*0.384*	*0.199*	*0.569*
27	**resp**	10.442	8.225	12.659	11.857	9.779	13.934	6.405	4.749	8.061	2.640	1.583	3.697	3.376	1.736	5.016
28	**sta**	0.208	0.148	0.341	0.211	0.148	0.337	0.217	0.126	0.393	0.129	0.079	0.177	0.110	0.053	0.167
29	**fat**	2.350	2.350	2.350	2.516	2.516	2.516	1.598	1.598	1.598	0.603	0.603	0.603	0.543	0.543	0.543
30	**pro1**	0.214	0.214	0.214	0.210	0.210	0.210	0.156	0.156	0.156	0.068	0.068	0.068	0.072	0.072	0.072
31	**pro2**	0.060	0.060	0.060	0.063	0.063	0.063	0.055	0.055	0.055	0.022	0.022	0.022	0.020	0.020	0.020
32	**pro3**	0.219	0.219	0.219	0.231	0.231	0.231	0.176	0.176	0.176	0.068	0.068	0.068	0.063	0.063	0.063
33	**pro4**	0.120	0.120	0.120	0.127	0.127	0.127	0.111	0.111	0.111	0.045	0.045	0.045	0.040	0.040	0.040
34	**pro5**	0.437	0.437	0.437	0.474	0.474	0.474	0.413	0.413	0.413	0.168	0.168	0.168	0.153	0.153	0.153
35	**pro6**	0.074	0.074	0.074	0.078	0.078	0.078	0.079	0.079	0.079	0.029	0.029	0.029	0.026	0.026	0.026
36	**pro7**	0.271	0.271	0.271	0.294	0.294	0.294	0.227	0.227	0.227	0.097	0.097	0.097	0.093	0.093	0.093
37	**pro8**	0.279	0.279	0.279	0.301	0.301	0.301	0.237	0.237	0.237	0.101	0.101	0.101	0.103	0.103	0.103
38	**pro9**	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
39	**glns**	0.737	0.620	0.844	0.789	0.666	0.888	0.686	0.632	0.746	0.262	0.223	0.300	0.275	0.253	0.287
40	**glus**	1.325	0.933	1.479	1.461	1.384	1.479	1.212	1.003	1.264	0.442	0.281	0.506	0.451	0.361	0.479

Table 4.5.1 Continued

Reaction	Enzymes	H/H	Confidence Interval (90%)		M/H	Confidence Interval (90%)		M/M	Confidence Interval (90%)		M/L	Confidence Interval (90%)		L/L	Confidence Interval (90%)	
#		Mean	Lower	High	Mean	Lower	High	Mean	Lower	High	Mean	Lower	High	Mean	Lower	High
41	**gdh**	1.193	-0.331	2.551	1.979	0.718	2.636	0.948	-0.150	2.017	0.476	0.069	0.884	0.573	0.221	0.824
42	**gad**	1.651	0.160	2.960	2.467	1.230	2.960	1.396	0.306	2.434	0.637	0.229	1.012	0.745	0.391	0.958
43	**gatf**	1.651	0.160	2.960	2.467	1.230	2.960	1.396	0.306	2.434	0.637	0.229	1.012	0.745	0.391	0.958
44	**gatb**	*0.533*	*0.039*	*0.999*	*0.656*	*0.146*	*0.999*	*0.558*	*0.049*	*0.999*	*0.509*	*0.000*	*0.999*	*0.497*	*0.000*	*0.999*
45	**gabd**	1.651	0.160	2.960	2.467	1.230	2.960	1.396	0.306	2.434	0.637	0.229	1.012	0.745	0.391	0.958
46	**icl**	0.017	0.000	0.049	0.003	0.000	0.010	0.001	0.000	0.005	0.001	0.000	0.005	0.002	0.000	0.009
47	**ms**	0.017	0.000	0.049	0.003	0.000	0.010	0.001	0.000	0.005	0.001	0.000	0.005	0.002	0.000	0.009
48	**pyrx**	0.476	0.153	0.740	0.505	0.200	0.740	0.433	0.168	0.632	0.107	0.000	0.220	0.079	0.000	0.191
49	**malx**	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
50	**scnx**	0.707	0.491	0.923	0.379	0.184	0.575	0.552	0.408	0.696	0.116	0.009	0.222	0.013	0.000	0.050
51	**glnx**	*0.979*	*0.939*	*0.999*	*0.983*	*0.963*	*0.999*	*0.976*	*0.919*	*0.999*	*0.945*	*0.792*	*0.999*	*0.962*	*0.894*	*0.999*
52	**c1f**	0.219	0.219	0.219	0.231	0.231	0.231	0.176	0.176	0.176	0.068	0.068	0.068	0.063	0.063	0.063
53	**x1b**	0.805	0.795	0.815	0.787	0.776	0.798	0.476	0.448	0.504	0.077	0.000	0.162	0.137	0.123	0.150
54	**glyx**	0.219	0.219	0.219	0.231	0.231	0.231	0.176	0.176	0.176	0.068	0.068	0.068	0.063	0.063	0.063
55	**c1x**	0.219	0.219	0.219	0.231	0.231	0.231	0.176	0.176	0.176	0.068	0.068	0.068	0.063	0.063	0.063
56	**glnsb**	1.459	1.436	1.479	1.456	1.431	1.479	1.244	1.223	1.264	0.494	0.484	0.505	0.469	0.459	0.479
57	**penx**	0.002	0.000	0.008	0.002	0.000	0.011	0.002	0.000	0.012	0.111	0.031	0.177	0.093	0.014	0.167
58	**frux**	0.001	0.000	0.007	0.001	0.000	0.005	0.002	0.000	0.007	0.006	0.001	0.010	0.006	0.001	0.010

where there is a partitioning of the flux into multiple branches of the network. The node under consideration can be either rigid or flexible. A node is said to be 'rigid' if the ratio of the carbon flow through the node remains the same irrespective of changes to the total flux coming into the node and is 'flexible' if the ratio of the carbon flow changes with a change in the incoming flux.

For rigid nodes, genetic modification will be necessary to alter the metabolic flow in the desired direction. However, in the case of flexible nodes, the channeling of the precursor metabolite can be modified by the system without the need of any major genetic modification [54] . For, example if we consider the node where the flux partitions into the pentose phosphate and glycolysis pathways, our statistical analysis (Figure 4.5.1) indicates that the flux ratios are significantly different for all five temperature conditions. This difference implies that the soybean system inherently takes care of the increased or decreased requirement of flux through the PPP or glycolysis branches of the reaction network. Hence, a genetic manipulation of this particular node is not required to change the carbon flow. Thus, in the case of the soybean system, as shown in Figure 4.5.1, our results indicate that the network is flexible, though no particular trend can be detected. However, a two-compartment model would be required before we can make the conclusion of the flexible network with more confidence.

The observations from section 4.4.2 about the GABA shunt and pyruvate effluxes for M/M temperature treatment were evident in the remaining four temperature treatments as well. For the lower temperature treatments, the anaplerotic reaction of PEP to OAA with respect to the pyruvate dehydrogenase reaction was negligibly small. However, for the higher temperature treatments, there was a small flux through the PEP to OAA reaction. The amount of protein accumulated was much greater in the higher temperature treatments as compared to the lower temperature treatments, with the M/H producing maximum protein. Figure 4.5.2 shows that the value of the ppcf/pdh flux ratio is the highest for the M/H, followed by the

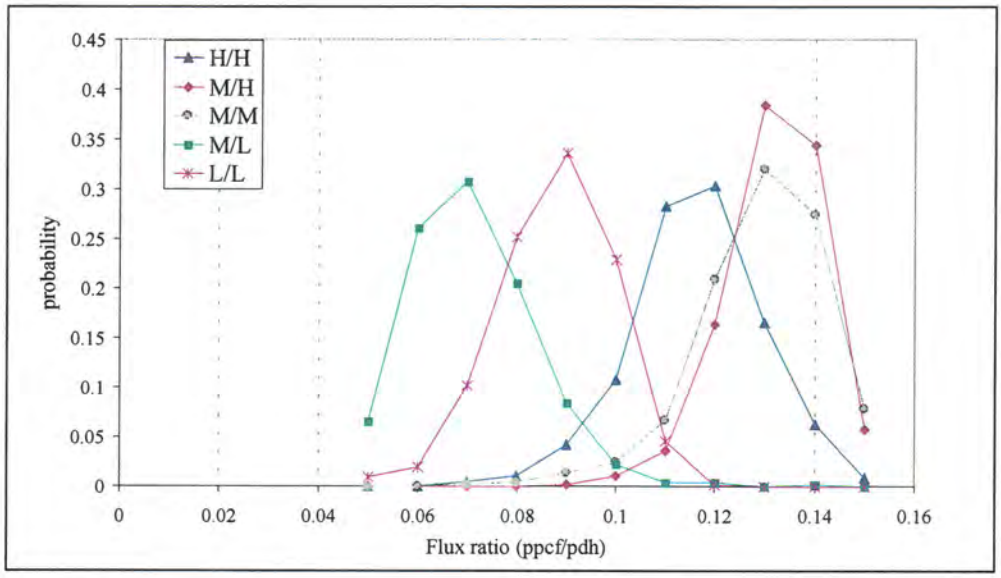
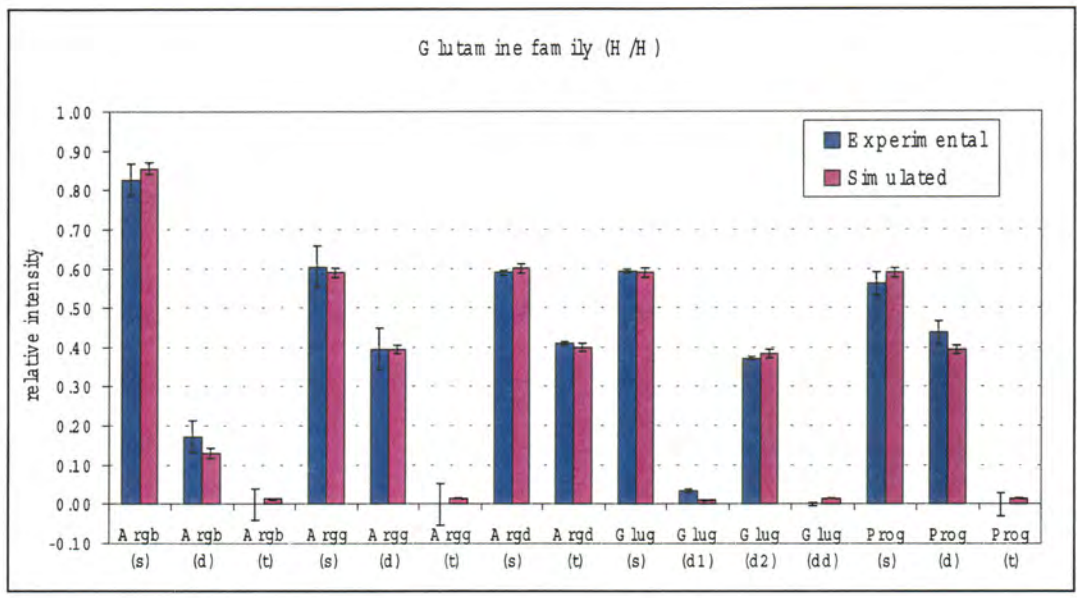


Figure 4.5.2 Probability distribution of flux ratio of PEP carboxylase to pyruvate dehydrogenase (ppcf/pdh) for the five different temperature treatments

M/M and H/H temperature treatment respectively. Thus, the flux through anaplerotic reactions do increase as required in the system. In our case the need for the precursor metabolite to produce more protein provides the incentive. For the remaining 4 temperature treatments also, the glyoxylate shunt was inactive.

As observed in the case of the M/M temperature treatment, the NMR intensity compared well (3-7 % difference) for Glu, OAA and Pyr family of amino acids but there were considerable differences (more than 10 %) for the PEP and P5P family of amino acids. Figure 4.5.3 represents the intensity comparison for the glutamine family of amino acids for the H/H and the PEP family of amino acids for the M/L as a representation of the comparison. The comparison of the input intensities for all the temperature treatments is discussed in the next section.

(a)



(b)

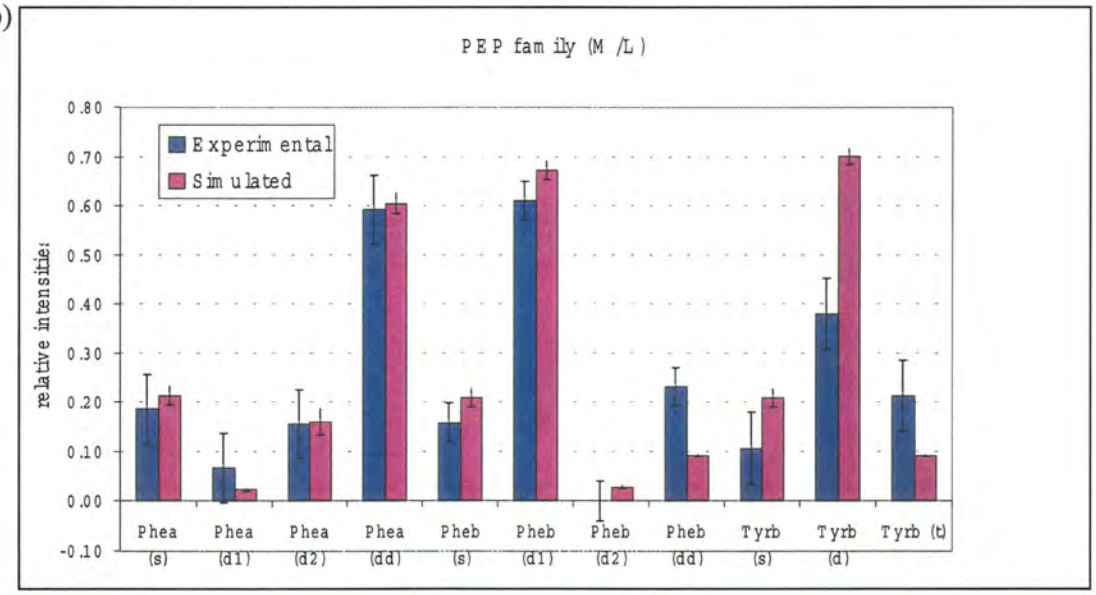


Figure 4.5.3 Comparison of experimental and simulated intensities for (a) Glutamine family of proteinogenic amino acids for H/H treatment and (b) Pyruvate family of proteinogenic amino acids for M/ L treatment (a, b, g and d indicate α, β, γ and δ carbon atom; s, d1, d2, dd and t indicates singlet, doublet, second doublet, double doublet and triplet respectively)

4.5.2 Comparison of input NMR intensities for the five temperature treatments

Consider the reaction network as shown in Figure 4.3.1. In the case of the pyruvate family of amino acids the δ^1 carbon of Leu, the β carbon of alanine and γ^1 carbon of Val reflect the same carbon atom of pyruvate respectively. Hence, the relative intensities should be the same for these carbon atoms in these three amino acids. As seen in Figure 4.5.4, for the M/M temperature treatment, the relative intensities of Ala and Val agree with each other, whereas the δ^1 carbon of Leu shows a 30 % difference from Ala and Val. This discrepancy is also noticed for the remaining temperature treatments. The δ^2 carbon of Leu and γ^2 carbon of Val agree with each other and reflect the carbon 3 of their precursor metabolite pyruvate respectively. Val and Leu are formed from the intermediate metabolite 2-Oxoisovalerate. Val is formed directly from 2-Oxoisovalerate whereas there are intermediate steps (Figure 4.5.5) in the formation of Leu. Thus, it is possible that the discrepancy between the δ^1 carbon of Leu with the γ^1 carbon of Val results from one or more of these three reactions being reversible leading to a carbon skeleton rearrangement as opposed to a compartmentation possibility. This possibility of reversibility needs to be investigated further.

There is a marked increase in the double doublet intensity of Alaα and the doublet intensity of Glyα for the lower treatments as compared to the M/M and higher temperature treatments. The increase in the double doublet intensity indicates that Alaα and Glyα are being formed from intact molecules of hexose phosphates. This increase in intensity is reflected as an increase in the flux through glycolysis as compared to PPP (Table 4.5.1).

Since PEP is metabolized directly to pyruvate, Alaα should show the same intensity fractions as Pheα. Both carbon atoms (Alaα and Pheα) are derived from second carbon of PEP respectively. It was found, however, for all the temperature treatments that the intensity fractions of Alaα and Pheα were different (Appendix E). When the 1D extract of the Pheα and Alaα were normalized to their respective maximum intensity and superimposed over

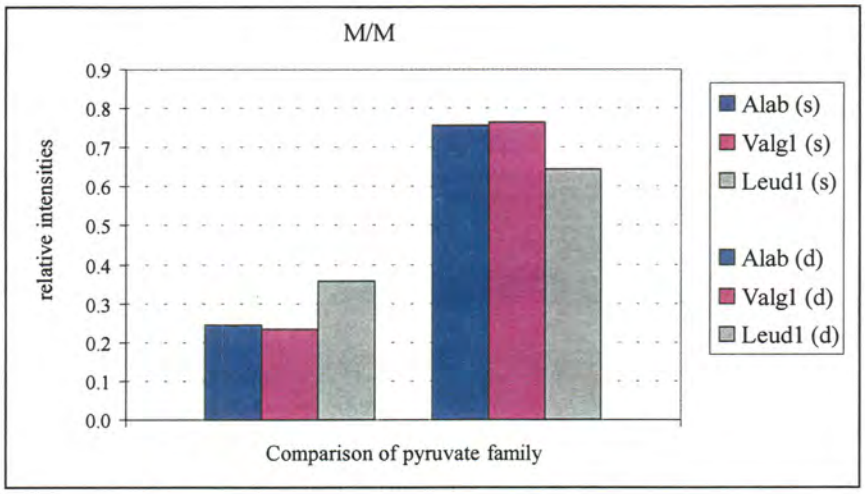


Figure 4.5.4 Comparison of relative intensities of δ^1 carbon of Leu, the β carbon of alanine and γ^1 carbon of valine for M/M temperature treatment (b, d and g indicate β, δ and γ carbon atom; s and d indicate singlet and doublet respectively)

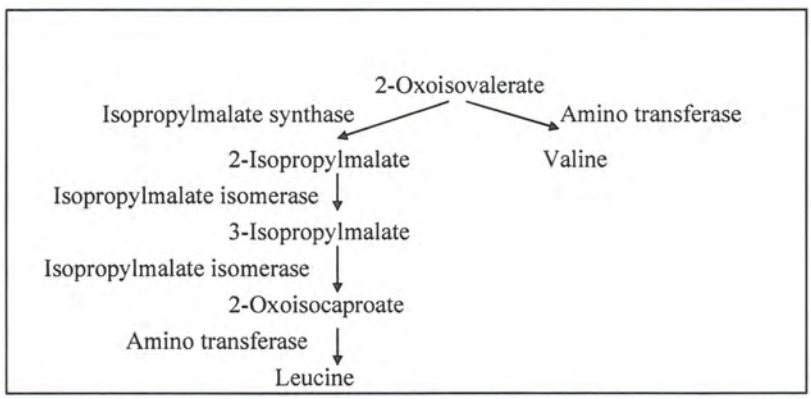


Figure 4.5.5 Biosynthesis of Leu and Val from 2-Oxoisovalerate

each other, their intensity patterns compared very well (Figure 4.5.6). This result suggests that the difference in the relative intensity fraction of Pheα and Alaα maybe an artifact of the spectral processing program.

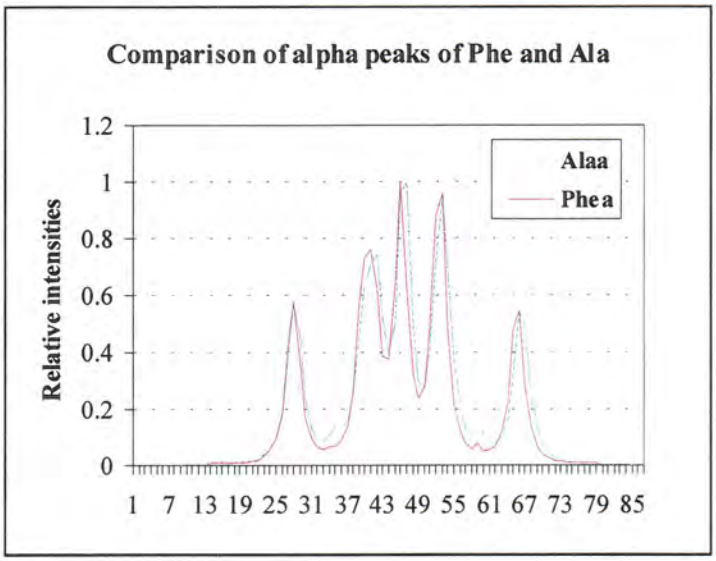


Figure 4.5.6 Comparison of relative intensities of NMR peaks for Alaα and Pheα data from the M/H treatment

4.5.3 Additional observations for the M/L and L/L treatments

To estimate the partitioning of the fluxes between the PPP and glycolysis accurately, five fluxes need to be determined. Transketolase (A and B) and transaldolase enzymes catalyzing the PPP reactions, the hexose isomerase reversibility and the entry flux into the PPP catalyzed by 6-phosphogluconate dehydrogenase enzyme. The independent measurements required for detecting the partitioning are obtained from the amino acids. His and Phe give 2 independent measurements each and glycine gives one measurement. Thus, we have 5 independent measurements to estimate 5 unknown variables.

In the case of the M/L and L/L temperature treatments, since the protein concentration was lower than that of the higher temperature treatments, the signal to noise ratio of the NMR spectrum was weak. In the case of the L/L, even though there were five independent measurements, the noise level was very high. For M/L temperature treatment, only one independent measurement from His was obtained. Thus, sufficient measurements were not available to estimate the branchpoint distribution. When the simulations were carried out without any additional constraints to the system, the PPP flux was found to be structurally not identifiable with a very high standard deviation.

From the H/H, M/H and M/M temperature treatments, the values of the means of the extent of reversibilities for the transketolase and transaldolase reactions were very high in the range of 0.9 to 0.999. The actual values are tabulated in Table 4.5.2. To compensate for the lack of measurements, the values of the extent of reversibilities for the transketolase and transaldolase reactions were set as shown in Table 4.5.2 for the M/L and L/L temperature treatments respectively.

The simulations output discussed in sections 4.5.1 and 4.5.2 are the result after setting the extent of reversibilities limit. The results of the simulation output for the lower treatments are discussed in detail in section 4.5.1. The main observation regarding the flux values for the lower treatments is that the confidence limits for the fluxes are very large. The reason for the larger confidence intervals is the high noise level in the spectrum for the lower temperature treatments. The M/L and L/L temperature treatments had approximately 5.4 mg of protein each (Table 4.2.1). As explained in section 4.3.4, to run a 2D HSQC experiment for 20 hours, the minimum amount of protein required is 7 mg when the system is fed with 10% ^{13}C labeled sucrose. Around 30 hour 2D HSQC measurements were carried out for the lower temperature treatments with double the number of scans to compensate for the reduced amounts of protein. However, the signal to noise ratio increases only 1.4 times when we double the scans and this is clearly not sufficient for a good resolution NMR spectrum.

Table 4.5.2 Extents of reversibility for the pentose phosphate pathway reactions

Temperature Treatment	Mean values of E →	E_{tktA}	E_{tal}	E_{tktB}
H/H	From simulation	0.969	**0.940**	0.999
M/H	From simulation	0.973	0.963	0.999
M/M	From simulation	**0.965**	0.941	0.999
M/L (set)	E_{min}	0.965	0.940	0.9989
	E_{max}	0.999	0.999	0.999
L/L (set)	E_{min}	0.965	0.940	0.9989
	E_{max}	0.999	0.999	0.999

E_{tktA}, E_{tktB}, E_{tal} indicate the extent of reversibility for transketolase A, tranketolase B and transaldolase enzyme. E_{min} and E_{max} indicate the lower and upper limits set for lower temperature treatments.

Hence, a minimum of 7 mg of protein is required to run 2D HSQC experiment.

In the case of the lower temperature treatments, the efflux of saccharides and efflux from P5P are very high (Table 4.5.1). The extent of reversibility for the G6P ↔ F6P reaction (catalyzed by hexose isomerase) was found to be much higher for the lower temperature treatments (with confidence interval of 0.1 to 0.9) as compared to the higher temperature treatments (with a confidence interval of 0-0.5). It is possible that the values set for the extent of reversibilites do not hold true for the lower temperature treatments. Alternatively, since the extent of reversibility is high for the hexose isomerase reaction, the flux going through the PPP is sent out as P5P efflux. Thus, flux through the hexose isomerase reaction must be quantified more accurately to determine flux partitioning at the branchpoint of PPP and glycolysis.

4.5.4 Identification of unknown peaks in the NMR spectra

Figure 4.3.2, which represents 2D HSQC spectrum of hydrolyzed protein sample indicates the presence of four peaks (labeled 1, 2, 3 and 4 in Figure 4.3.2) whose chemical shift could not be assigned to any amino acid. Experiments with hydrolyzed glucose and sucrose samples indicated that these peaks are a result of hydrolyzed sugar molecules from glycosylated protein. The monomer attached to the proteins is generally mannose. The precursor for sugar for glycosylation is most probably from the hexose phosphates (G6P and F6P) from the glycolysis pathway. These peaks are relatively strong as compared to the amino acids in the spectrum and could possibly help in the prediction of the PPP to glycolysis flux ratio more accurately.

There are starch samples extracted from the biomass for the five temperature treatments. Preliminary analysis (2D HSQC) has been carried out on the hydrolyzed starch sample for the M/M temperature treatment. The isotopomer distributions of the gxx (g represents glucose and xx represents unknown) peaks (1, 2, 3 and 4 in Figure 4.3.2) obtained from the hydrolyzed protein are different from those obtained from the starch sample. Starch is synthesized in the plastids exclusively [55]. The difference in the isotopomer distribution of gxx peaks could mainly be due to the fact that starch sample reflects the plastidic information and the information from the protein spectrum reflects the cytosolic information. The information from the 2 sets of gxx peaks could be used to construct a two-compartment model for central carbon metabolism in soybeans.

5. CONCLUSIONS

A single compartmental model comprising all the known reactions in the central carbon metabolism of soybean was constructed. A combination of biosynthetically directed bond labeling experiment and 2D HSQC technique was used to quantify the intracellular fluxes in the metabolic pathway of soybean. The model was used to study the temperature effects on the metabolism of the developing soybean cotyledon. The model estimated the metabolic fluxes with 90 % confidence limits for the flux values.

Metabolic flux analysis (MFA) of the developing soybean cotyledon enabled an understanding of the central carbon metabolism in soybean cotyledons. Important observations on pathway functioning were made and potential targets for genetic targets could also be identified. For all the temperature treatments analyzed, there was a dilution in labeling information in the lower part of the reaction network including the reactions of tricarboxylic acid (TCA) cycle, glyoxylate and γ-aminobutyrate (GABA) shunt and the glutamine metabolism. The dilution was a result of the unlabeled carbon from glutamine, which was supplied as a nitrogen source. As a result, the GABA shunt could not be identified clearly. The problem of non-identifiability can be solved using labeled glutamine to increase the labeling information in the TCA cycle and glutamine metabolism. The glyoxylate shunt was found to be inactive for all temperature treatments. The organic acid effluxes, especially pyruvate were also not statistically identifiable. The complex relationship between the NMR data and the metabolic fluxes for the pyruvate part of the network, reflected as large confidence intervals in the estimation of pyruvate efflux. It was concluded that extracellular measurements like organic acid effluxes, left over substrate were important measurements for estimating the fluxes more accurately. The maximum limit for the extent of reversibility (E) for a reversible reaction was set at 0.999. A value of 0.9999 for E was impractical and a value of 0.99 did not estimate the optimum correctly.

The flux through the PPP reaction was found to be 30 % of the flux through the glycolysis pathway for the M/M temperature treatment. The probability density function, which was plotted for the flux ratio of PPP to glycolysis was significantly different between all the five temperature treatments except in the case of the M/M and the M/L temperature treatment. The difference suggests that the PPP to glycolysis branchpoint is a 'flexible' node and not a 'rigid' node. From the definition of 'rigid' and 'flexible' nodes, it is clear that genetic modification will be necessary to alter the metabolic flow in the desired direction in the case of rigid node whereas flexible nodes don't need of any major genetic modification. Thus, in the case of the soybean system, our results indicate that the network is flexible regards carbon partitioning, though no particular trend could be detected. However, a two-compartment model would be required before the conclusion of the flexible network is made with more confidence.

The anaplerotic reactions were observed to be higher for the H/H, M/H and M/M temperature treatment as compared to the M/L and L/L temperature treatment. The higher values of the anaplerotic reactions can be explained by the need for an increase in the precursor metabolites to compensate for the increased protein quantities in the higher temperature treatments. The ratio of the pyruvate dehydrogenase reaction to the PEP carboxylase reaction was around 8-10 for the five treatments indicating that the most of the flux channels into lipids and the TCA cycle. Thus, down regulating the PEP carboxylase gene in soybean would not help in increasing the carbon partitioning, say in lipids as the flux through PEP carboxylase is very small as compared to flux through pyruvate dehydrogenase.

In the case of the lower temperature treatments, the amount of protein accumulated *in vitro* was approximately 5.4 mg. The resultant NMR spectrum had a very high noise signal and the model could not identify the flux distribution between the PPP and glycolysis reactions due to fewer required number of independent measurements. There was no net flux through the pentose phosphate reactions for the higher temperature treatments. An

assumption of zero net flux in the PPP reactions was made for the lower temperature treatments by setting values of E from the higher temperature treatments and the fluxes were evaluated. However, for the lower temperature treatments, the hexose isomerase reversibility was very high as compared to the higher temperature treatments. The penx and sac effluxes were found to be very high as a result of the increased reversibility of the hexose isomerase reactions for the lower temperature treatments. A minimum of 7 mg of protein (or more), hence, is required to procure a good resolution NMR spectrum for the soybean system, when fed with ^{13}C labeled sucrose.

On comparing the simulated and the experimental intensities, it was observed that the program estimated the intensities for the amino acids in the OAA, Glu and Pyr family of amino acids well. However, in the case of PEP and P5P family of amino acids the difference between simulated and the experimental intensities was found to be more than 10%. A considerable difference between the input intensities of Pheα and Alaα was observed. However, it was realized that the difference was more an artifact of the spectral processing program, rather than the suggestion of compartmentation. The intensity of the Leuδ^1 peak did not agree with the Val γ^1 and Alaβ peak, although they reflect the same carbon atom of the precursor metabolite. However, Leuδ^2 agreed with the Val γ^2 peak very well. There are three intermediate reactions in the biosynthesis of leucine from 2-Oxoisovalerate to leucine. Thus, the discrepancy between the Leuδ^1 and Val γ^1 peak may be a result of the likelihood that one or more of these reactions in the biosynthesis of leucine are reversible.

The prominent unknown peaks noted in the HSQC spectrum (for all the five temperature treatments) were inferred from preliminary experiments to be products of hydrolyzed sugar molecules from the glycosylated proteins. The monomer attached to the proteins is generally mannose. The precursor for sugar for glycosylation is most probably from the hexose phosphate in the glycolysis pathway. The isotopomer distribution information obtained from these peaks was different from the distribution obtained from the

hydrolyzed starch sample for the M/M temperature treatment. The isotopomer distribution obtained from both sources can provide important information about compartmentation as starch is exclusively synthesized in the plastid.

Future Direction

The difference in the isotopomer distribution from the hydrolyzed starch and hydrolyzed hexose phosphate molecules lead to the likely possibility of compartmentation in the pentose phosphate pathway and glycolysis part of carbon metabolism of soybean. To explore this, the next step would be to add the required reactions and the carbon skeleton rearrangement to the mathematical model. The hydrolysis of starch for all the five temperature treatments needs to be carried out and a 2D HSQC experiment has to be performed and analyzed for the five temperature treatments. The aromatic peaks of His, Phe and Tyr needs to be analyzed for the two-compartment model for more independent measurements. 2D HSQC experiments also need to be carried out for the five temperature treatments with a new sweep width region of the NMR spectrum to get the aromatic peaks. Using the isotopomer data from both the compartments, we can quantify the fluxes in the central carbon metabolism of soybean. The reactions in the biosynthesis of Leu need to be studied in detail to account for the discrepancy between the intensities of Leuδ^1 and Val γ^1, thus improving the existing model. Finally, a two compartment model would surely give us a better insight of the working of the soybean system.

APPENDIX A: EXTRACTION PROTOCOLS FOR LIPIDS, PROTEIN AND STARCH

i. Hexane extraction of neutral lipids

Sample weight should be in the range of 50 to 100 mg while using microfuge tubes and the heating blocks unit. We used freeze-dried, ground embryo tissue (seed coat removed prior to freeze-drying). The protocol used is as follows:

1. Label and preweigh disposable, borosilicate glass tubes (5 times number of samples for 5 sets of hexane washes)
   when air drying or using hot water bath, use 13x75 mm tubes
   when drying with heat blocks, use 10x75 mm tubes
2. Preheat heating block
3. Weigh sample and place in microfuge tube
4. Add 1 ml high grade hexane to each microfuge tube and vortex
5. Place tubes in the heating block at 40°C for about an hour
6. Centrifuge 1 min on benchtop centrifuge at high speed
7. Remove supernatant (leaving a small amount in the microfuge tube) to preweighed glass tubes and keep under hood
8. Repeat 4 times from step 4 removing supernatant to additional sets of preweighed tubes each time
9. Dry pellets in heating block (turned off but cooling over night) or speed vacuum for 15 minutes
10. Store defatted pellet at -20°C
11. Weigh tubes when they have come to room temperature and dryness (about 48 hours), subtract the preweight and sum the differences to determine the total weight of oil extracted
12. Determine percent of original sample weight. Correct to 13% moisture content. Assume freeze-dried tissue is 0% moisture content

Formula for calculations

% moisture content correction:

$$\frac{(1\text{-moisture content}_{initial}) \times \text{tissue weight}_{initial}}{(1\text{-moisture content}_{final})} = \text{tissue weight}_{final}$$

% oil calculation:

$$\frac{\text{mg oil}}{(\text{tissue weight}_{final})\ \text{mg}} \times 100 = \%\ \text{oil}_{final}$$

ii. Protein extraction

Procedure for Kjeldahl digestion

1. Add slightly less than 1 scoop of catalyst to each tube
2. Add one increment (3.5 ml) of sulfuric acid to each tube
3. Take the rack of tubes for digestion
4. Program block to digest at 382°C for 1 hour
5. Turn on hood fan and carefully place rack on Block
6. Place rack "ends" on the rack
7. Carefully place manifold on rack
8. Hook up black suction tube to back of manifold and turn on water
9. Place hood sash approximately 4 inches from the bottom

When digestion is complete

1. Turn digestion timer off
2. Fill beaker ¼ to 1/3 with water for rinsing suction tube
3. Remove "rack ends" from rack
4. Remove black tube from manifold and rinse by sucking water
5. Carefully lift off manifold and turn over
6. Carry manifold upside down to sink for rinsing
7. Place manifold on manifold holder
8. Carefully lift rack of tubes from Block and set in hood
9. Lower sash to within 3 inches from bottom
10. Leave fan on and let tubes cool for approximately 10 minutes

Procedure for Kjeldahl distillation

Preparation

1. When tubes are cool, add water (15 ml) to each tube and let cool
2. Turn on boiling flask heaters
3. Add distilled water to boiling flasks to slightly below mark
4. Turn on condensation water
5. Add one "squirt" (5 ml) indicator solution to each of 40 flasks

6. Fill sodium hydroxide column (Kjeldahl tubes must be cooled down to air temperature before distillation process can begin)
7. After water has been boiling for a few minutes, distillation can begin

Distillation

1. Slowly add 1 increment of sodium hydroxide to 2 Kjeldahl tubes
2. "Shut off" boiling valve
3. Remove distilled Kjeldahl tube which is on apparatus
4. Place new Kjeldahl tube firmly on distillation apparatus
5. Turn boiling valve back on
6. Place indicator flask under condensation tube
7. After collecting 15-20 ml of distillate, remove indicator flask
8. Repeat steps 1 through 7
9. Dump contents from distilled Kjeldahl tubes into Kjeldahl waste container
10. Rinse tubes thoroughly and place them back on the rack
11. When finished, carefully attach rack holder to top of rack
12. Dump excess water from tubes in the sink
13. Place upside down in oven to dry

iii. Procedure for starch measurement

Starch in the insoluble residue was determined with an enzymatic hydrolysis method by Microsample Plate Reader

1. 20-80 mg of dried defatted powder was weighed and 1 ml of 80 % ethanol was added. This was kept in a heater at 60°C for 20 min. It was then centrifuged at 5500Xg for 10 min and the supernatant was removed. The procedure was repeated four times with 1 ml of 80 % ethanol.
2. Supernatant was collected in glass tube, dried on heater or water both at 60°C. The dried supernatant in glass tube was sealed with polyfilm, stocked in freezer at –20°C. It can be used to measure soluble sugar.
3. The pellet was transferred to glass tube with 1.5-7.5 ml water, and covered by aluminum foil and autoclaved at 131°C for 1 h.
4. 10 mg of standard starch was weighed and 1.5 ml water was added to it and autoclaved along with sample.
5. Gelatinized starch was digested by adding 1.5-7.5 ml of 0.1 M citrate buffer (pH 5.0), containing 2 mg amyloglucodase and incubated overnight in waterbath at 30 °C.
   For example, if 15 mg sample (DW) was taken and 1.5 ml water was added to it and autoclaved. Then 1 .5 ml of 0.1 M citrate buffer was added containing 0.375 mg

amyloglucodase. Hence, ratio of Amyloglucodase mg/ sample mg is calculated as 0.375 mg / 15 mg = 0.025 mg/mg.DW

0.1M citrate buffer [pH 5.0] is prepared as:

A: 0.1 M solution of citrate acid (21.01 g in 1000 ml)

B: 0.1 M solution of sodium citrate (29.41 g $C_6H_5O_7Na_3$. $2H_2O$ in 1000ml) 20.5 ml of A and 29.5 ml of B, diluted to a total of 100 ml, pH 5.0

6. The part of digested aliquot was centrifuged at 13,000Xg for 10 min. Glucose in the supernatant was determined by glucose oxidase, using a Microsample Plate Reader. Chemical solutions included:

- Tris-phosphte-glycerol buffer: 36.3 g of Tris , 56.5g of sodium dihydrogen phosphate dihydrate ($NaH_2PO_4.2H_2O$) or 49.9836 g ($NaH_2PO_4.H_2O$) dissolved in 400 ml of water. Plus 400 ml of glycerol and then diluted to 1 liter with water.
- Reagent A consisted of 30 mg of glucose oxidase, 3 mg of peroxidase dissolved in 50 ml of Tris-phosphate-glycerol buffer.
- Reagent B consisted of 10 mg of o-dianisidine dissolved in 50 ml of the above described buffer. Working Reagent consisted of equal volumes of Reagent A and Reagent B, prepared fresh daily.

7. The procedure involved the addition of 50 µl of sample (containing 10-1000ug of D-glucose/ml) to 200 µl of working glucose oxidase reagent in a microsample plate.
8. The plate, with sample, was incubated at 37°C for 30 min. 50 µl of concentrated HCl was added and the absorbance was measured at 492 nm.
9. D-a-glucose standard curve:

   Prepared 0.25 µg/µl solution of D-a glucose.

   a) 0, 2, 4, 6, 8, 10, 12, 14, 18, 20 µl of 0.25µg /µl D-a glucose was taken into each plate well, respectively,
   b) 50, 48, 46, 44, 42 40, 38, 36, 34, 32, 30 µl water was added to above plate well.
   c) 200 µl of gluxose oxidase working reagent was added in a microsample plate well. The plate was mixed at slow speed for 30 s on a vortex and then incubated at 37°C for 30 min. 50 µl of concentrated HCl was added to each well and the plate was mixed at slow speed for 10 s on a vortex mixer. The absorbance was measured at 492 nm.
   d)

$$\text{Starch content (mg / DW mg)} = \frac{\text{slope of standard glucose* Absorbance of sample } \mathbf{0.9} \text{ diluted time}}{\text{Sample weight mg (DW)}}$$

Table B.1: HPLC analysis of proteinogenic amino acids for five temperature treatments (T-test carried out to indicate whether statistically different from M/M temperature treatment)

Amino acid	pMoles H/H	Mole %	pMoles M/H	Mole %	pMoles M/M	Mole %	pMoles M/L	Mole %	pMoles L/L	Mole %	Statistically different from M/M at P=0.05 (Yes (Y) or No (N)) H/H	M/H	M/L	L/L
Ala	574	7.43	610	8	341	6.95	493	7.94	463	7.84	N	Y	Y	Y
Asx	941	12.2	973	12.8	547	11.1	626	10.1	683	11.5	N	Y	N	N
Glx	1176	15.2	1272	16.7	791	16.1	986	15.9	1049	17.7	N	N	N	N
Phe	292	3.78	283	3.71	195	3.97	241	3.88	210	3.55	N	N	N	N
Gly	597	7.73	546	7.17	372	7.58	450	7.24	419	7.09	N	N	N	N
His	195	2.53	190	2.49	104	2.12	147	2.37	167	2.83	Y	Y	N	Y
Ile	363	4.71	349	4.59	230	4.68	297	4.79	260	4.4	N	N	N	N
Lys	477	6.18	483	6.34	317	6.46	422	6.8	376	6.36	N	N	N	N
Leu	594	7.69	579	7.6	466	9.49	512	8.25	448	7.57	Y	Y	Y	Y
Met	87	1.12	70	0.92	80	1.63	117	1.89	50	0.85	Y	Y	Y	Y
Pro	541	7	415	5.45	293	5.96	333	5.36	309	5.22	Y	Y	N	Y
Arg	536	6.94	536	7.04	311	6.34	482	7.77	451	7.63	N	N	Y	Y
Ser	380	4.92	385	5.06	271	5.53	306	4.93	294	4.98	Y	N	N	N
Thr	320	4.14	294	3.86	157	3.2	265	4.27	251	4.26	Y	Y	Y	Y
Val	461	5.97	449	5.9	304	6.2	377	6.07	337	5.71	N	N	N	N
Tyr	190	2.46	185	2.43	130	2.66	156	2.51	142	2.41	N	N	N	N
Total	7727	100	7621	100	4915	100	6215	100	5915	100				

APPENDIX C: REACTION NETWORK

Table C.1 List of reactions representing the central carbon metabolism of developing soybean cotyledon

Reaction #	Reaction name/Enzymes	Reaction
1	ini1	Sucrose → G6P + F6P
2	ini2, f	Glu → Glu
3	hxi, f	G6P → F6P
4	hxi, b	F6P → G6P
5	gnd	G6P → P5P + CO_2
6	tktA, f	P5P + P5P → S7P + T3P
7	tktA, b	S7P + T3P → P5P + P5P
8	tal, f	S7P + T3P → F6P + E4P
9	tal, b	F6P + E4P → S7P + T3P
10	tktB, f	P5P + E4P → F6P + T3P
11	tktB, b	F6P + T3P → P5P + E4P
12	pfk	F6P → T3P + T3P
13	gap	T3P → PEP
14	pyk	PEP → Pyr
15	pdh	Pyr → ACA + CO_2
16	cit	ACA + OAA → ∝KG + CO_2
17	suc1	∝KG → Scn + CO_2
18	fum1, f	Scn → Mal
19	fum1, b	Mal → Scn
20	fum2	Scn → Mal
21	mdh, f	Mal → OAA
22	mdh, b	OAA → Mal
23	ppc, f	PEP + CO_2 → OAA
24	ppc, b	OAA → PEP + CO_2
25	me, f	Pyr + CO_2 → Mal
26	me, b	Mal → Pyr + CO_2
27	resp	CO_2 → Efflux
28	sac	G6P → Efflux

Table 3.1 Continued

29	fat	ACA	→	Efflux to lipids
30	pro1	P5P	→	Efflux to amino acids
31	pro2	E4P	→	Efflux to amino acids
32	pro3	T3P1	→	Efflux to amino acids
33	pro4	PEP	→	Efflux to amino acids
34	pro5	Pyr	→	Efflux to amino acids
35	pro6	ACA	→	Efflux to amino acids
36	pro7	OAA	→	Efflux to amino acids
37	pro8	Glu	→	Efflux to amino acids
38	pro9	Gln	→	Efflux to amino acids
39	glns, f	Gln	→	Glu
40	glus	Glu	→	∝KG
41	gludh	∝KG	→	Glu
42	gad	Glu	→	GABA + CO_2
43	gat, f	GABA	→	SSA
44	gat, b	SSA	→	GABA
45	gabd1	SSA	→	Scn
46	icl	Scn	→	SSA
47	ms	ACA + GOX	→	Mal
48	pyrx	Pyr	→	Efflux
49	malx	Mal	→	Efflux
50	scnx	Scn	→	Efflux
51	in2, b	Glu	→	Glu
52	c1, f	T3P2	→	Gly0 + C1
53	c2, b	Gly0 + C1	→	T3P2
54	glyx	Gly0	→	Efflux
55	clx	C1	→	Efflux
56	glns, b	Glu	→	Gln
57	penx	P5P	→	Efflux
58	frux	F6P	→	Efflux

APPENDIX D: SUCROSE CONSUMPTION RATES USING CONSTANT YIELD AND Q_{10} = 2 METHODS

Table D.1 Calculation of carbon content in biomass for the H/H temperature condition

Amino acids	weight %	Mol wt	Protein mg	AA mg	AA moles	# of C	C moles	C moles/day
Ala	4.736	89.09	13.092	0.620	6.96E-06	3	2.09E-05	3.48E-06
Asx	11.341	133.1	13.092	1.485	1.12E-05	4	4.46E-05	7.44E-06
Glx	18.062	146.64	13.092	2.365	1.61E-05	5	8.06E-05	1.34E-05
Phe	5.013	165.19	13.092	0.656	3.97E-06	9	3.58E-05	5.96E-06
Gly	4.352	75.07	13.092	0.570	7.59E-06	2	1.52E-05	2.53E-06
His	2.515	155.16	13.092	0.329	2.12E-06	5	1.06E-05	1.77E-06
Ile	4.702	131.17	13.092	0.616	4.69E-06	6	2.82E-05	4.69E-06
Lys	3.243	131.17	13.092	0.425	3.24E-06	6	1.94E-05	3.24E-06
Lys	3.243	131.17	13.092	0.425	3.24E-06	6	1.94E-05	3.24E-06
Leu	10.621	146.19	13.092	1.390	9.51E-06	6	5.71E-05	9.51E-06
Met	1.866	149.21	13.092	0.244	1.64E-06	5	8.19E-06	1.36E-06
Pro	5.249	115.13	13.092	0.687	5.97E-06	5	2.98E-05	4.97E-06
Arg	8.452	174.2	13.092	1.106	6.35E-06	5	3.18E-05	5.29E-06
Ser	4.444	105.09	13.092	0.582	5.54E-06	3	1.66E-05	2.77E-06
Thr	2.920	119.12	13.092	0.382	3.21E-06	4	1.28E-05	2.14E-06
Val	5.556	117.15	13.092	0.727	6.21E-06	5	3.10E-05	5.17E-06
Tyr	3.687	181.19	13.092	0.483	2.66E-06	9	2.40E-05	4.00E-06
							TOTAL	**8.10E-05**

Table D.1 Continued

Lipids mg	Lipids moles	Lipids moles/day	T3P moles/day	A-CoA moles/day	C in T3P moles/day	C in A-CoA moles/day	Total C moles/day
6.4428	7.98E-06	1.33E-06	1.33E-06	3.19E-05	3.99E-06	6.38E-05	**6.78E-05**

Starch mg	Glucose mg	Glucose moles	Glucose moles/day	# of C	C moles/day
2.602	2.888	1.60E-05	2.67E-06	6	**1.60E-05**

Carbohydrates = Starch+ cell wall, Cell wall = 30%

Cellulose %	Hemi %	DW mg/day	Cellulose mg/day	Hemi mg/day	Mol wt of cellulose C6H12O6	Mol wt of hemi C6H10O5	Cellulose moles/day	Hemi moles/day	C in cellulose moles/day	C in hemi moles/day
15	15	6.883	1.033	1.033	180	162	5.74E-06	6.37E-06	**3.44E-05**	**3.82E-05**

Hemi indicates hemicellulose

Table D.2 Calculation of sucrose consumption by soybean embryos using constant yield method. Protein, starch and oil measured after 6 days in culture

Temp. Cond	DW Initial (mg)	DW Final (mg)	DW gain (mg)	Protein % of DW	Starch % of DW	Oil % of DW	Carb % of DW	Protein gain mg	Oil gain mg	Starch gain mg	Carb gain mg	k2/k1 factor *
H/H	10.10	66.90	56.80	26.30	5.20	16.67	35.20	14.94	9.479	2.95	19.99	1.38
M/H	9.80	65.00	55.20	29.30	5.50	18.40	35.50	16.17	10.16	3.04	19.60	1.34
M/M	**9.30**	**50.60**	**41.30**	**31.70**	**6.30**	**15.60**	**36.30**	**13.09**	**6.44**	**2.60**	**14.99**	**1.00**
M/L	10.20	28.10	17.90	30.70	7.80	13.60	37.80	5.49	2.43	1.40	6.77	2.31
L/L	8.60	27.50	18.90	28.20	5.70	11.60	35.70	5.33	2.19	1.08	6.75	2.19

Temp. Cond	Conversion factors ** Protein (g of glucose/ g of biomass constituent)	Oil	Carb	Glucose consumed for Protein mg	Oil mg	Carb mg	Starch mg	Total Glucose consumed w cw gm	wo cw gm	w cw moles	wo cw moles	Total Sucrose consumed w cw moles	wo cw moles
H/H													
M/H													
M/M	**1.793**	**3.030**	**1.211**	**23.47**	**19.52**	**18.16**	**3.15**	**0.06**	**0.05**	**3.4E-04**	**2.6E-04**	**1.7E-04**	**1.3E-04**
M/L													
L/L													

w cw indicates with cellwall, wo cw indicates without cellwall

*k2/k1 is the scaling factor for remaining temperature treatments with respect to M/M temperature treatment

*** conversion factors from literature [50]

Table D.2 Continued

Temp. Cond	Total Sucrose w cw	Total Sucrose wo cw	C in sucrose w cw	C in sucrose wo cw	DW gain mg/day	C moles *** (C from sucrose into) Protein	Oil	Starch	Cellwall	Carb	Total C into biomass	C lost as CO_2 w cw	C lost as CO_2 wo cw
	moles/day					moles/day						moles/day	
H/H					9.46								
M/H					9.20								
M/M	**2.8E-5**	**2.1E-5**	**3.4E-4**	**2.6E-4**	**6.88**	**8.1E-5**	**6.8E-5**	**1.6E-5**	**7.3E-5**	**8.9E-5**	**2.4E-4**	**1.0E-4**	**1.9E-5**
M/L					2.98								
L/L					3.15								

	CO_2 w cw m/day	CO_2 wo cw m/day	k2/k1 *	CO_2 w cw m/day	CO_2 wo cw m/day	C in CO_2 w cw m/day	C in CO_2 wo cw m/day	C moles **** Protein m/day	Oil m/day	Starch m/day	Cellwall m/day	Carb m/day	Total C in biomass m/day
H/H			1.38	1.4E-4	2.6E-5	1.4E-4	2.6E-5	9.2E-5	9.9E-5	1.8E-5	9.9E-5	1.2E-4	3.1E-4
M/H			1.34	1.4E-4	2.5E-5	1.4E-4	2.5E-5	9.9E-5	1.1E-4	1.9E-5	9.7E-5	1.2E-4	3.2E-4
M/M	**1.0E-4**	**1.9E-5**	**1.00**	**1.0E-4**	**1.9E-5**	**1.0E-4**	**1.9E-5**	**8.1E-5**	**6.8E-5**	**1.6E-5**	**7.3E-5**	**8.9E-5**	**2.4E-4**
M/L			2.31	4.4E-5	8.2E-6	4.4E-5	8.2E-6	3.4E-5	2.6E-5	8.6E-6	2.9E-5	3.8E-5	9.8E-5
L/L			2.19	4.7E-5	8.6E-6	4.7E-5	8.6E-6	3.2E-5	2.3E-5	6.6E-6	3.2E-5	3.8E-5	9.4-5

	C in suc w cw m/day	C in suc wo cw m/day	Total Sucrose consumed w cw m/day	Total Sucrose consumed wo cw m/day	Total Sucrose consumed w cw mg/day	Total Sucrose consumed wo cw mg/day
H/H	4.50E-4	3.36E-4	3.75E-5	2.80E-5	12.831	9.564
M/H	4.58E-4	3.47E-4	3.82E-5	2.89E-5	13.060	9.885
M/M	**3.40E-4**	**2.56E-4**	**2.83E-5**	**2.14E-5**	**9.682**	**7.307**
M/L	1.42E-4	1.06E-4	1.18E-5	8.83E-6	4.051	3.021
L/L	1.40E-4	1.02E-4	1.17E-5	8.52E-6	4.000	2.912

Table D.3 Calculation of sucrose consumption by soybean embryos using $Q_{10} = 2$ method

Temp. Cond	Total Sucrose w cw	wo cw	C in sucrose w cw	wo cw	DW	C moles *** (C from sucrose into) Protein	Oil	Starch	Cellwall	Carb	Total C into biomass	C in CO_2 w cw	wo cw
	moles/day		moles/day		mg/day	moles/day						moles/day	
H/H					9.47								
M/H					9.20								
M/M	**2.8E-5**	**2.1E-5**	**3.4E-4**	**2.6E-4**	**6.88**	**8.1E-5**	**6.8E-5**	**1.6E-5**	**7.3E-5**	**8.9E-5**	**2.4E-4**	**1.0E-4**	**1.9E-5**
M/L					2.98								
L/L					3.15								

	Temp °C	Initial DW mg	Final DW mg	DW gain Mg	Q10	T2-T1	T2-T1/10	2 exp (t2-t2/10)
H/H	35	10.1	66.9	56.8	2	8	0.8	1.741
M/H	35	9.8	62.9	53.1	2	8	0.8	1.741
M/M	27	9.3	54.2	44.9	2	0		
M/L	20	10.2	26.6	16.4	2	7	0.7	1.625
L/L	20	8.6	28.7	20.1	2	7	0.7	1.625

w cw indicates with cellwall, wo cw indicates without cellwall

*** conversion factors from literature [50]

Table D.3 Continued

	CO2 w cw moles/day	CO2 wo cw moles/day	Q10	CO2 w cw moles/day	C02 wo cw moles/day	**C moles** Protein moles/day	Oil moles/day	Starch moles/day	Cellwall moles/day	Carb moles/day
H/H			1.741	1.78E-04	3.28E-05	9.18E-05	9.97E-05	1.82E-05	9.99E-05	1.18E-04
M/H			1.741	1.78E-04	3.28E-05	9.89E-05	1.07E-04	1.87E-05	9.71E-05	1.16E-04
M/M	**1.02E-04**	**1.88E-05**		**1.02E-04**	**1.88E-05**	**8.10E-05**	**6.78E-05**	**1.60E-05**	**7.27E-05**	**8.87E-05**
M/L			1.624	6.29E-05	1.16E-05	3.38E-05	2.56E-05	8.61E-06	2.98E-05	3.84E-05
L/L			1.624	6.29E-05	1.16E-05	3.24E-05	2.31E-05	6.64E-06	3.14E-05	3.81E-05

Temp. cond	Total C in biomass moles/day	C in suc wcellwall moles/day	C in suc wocellwall moles/day	Total Sucrose consumed wcellwall moles/day	wocellwall moles/day
H/H	3.10E-04	4.88E-04	3.42E-04	4.063E-05	2.854E-05
M/H	3.22E-04	5.00E-04	3.54E-04	4.163E-05	2.954E-05
M/M	**2.38E-04**	**3.40E-04**	**2.56E-04**	**2.831E-05**	**2.136E-05**
M/L	9.79E-05	1.61E-04	1.09E-04	1.340E-05	9.121E-06
L/L	9.36E-05	1.56E-04	1.05E-04	1.304E-05	8.764E-06

Table E.1 Metabolic input intensities for carbon atoms of amino acids for the five temperature treatments

Precursor metabolite	Nucleus	Input Intensity H/H	Noise in spectrum H/H	Input Intensity M/H	Noise in spectrum M/H	Input Intensity M/M	Noise in spectrum M/M	Input Intensity M/L	Noise in spectrum M/L	Input Intensity L/L	Noise in spectrum L/L
T3P	Glyα (enr)	0.0940	0.0100	0.0700	0.0100	0.0729	0.0100	0.0750	0.0100	0.0906	0.0100
	Glyα (s)	0.4726	0.0015	0.5234	0.0013	0.4582	0.0015	0.3749	0.0156	0.3668	0.0029
	Glyα (d)	0.5274	0.0015	0.4766	0.0013	0.5418	0.0015	0.6251	0.0156	0.6332	0.0029
	Serα (enr)	0.0363	0.0100								
	Serα (s)	0.3126	0.0050	0.3444	0.0042	0.2720	0.0077			0.2361	0.0200
	Serα (d1)	0.2591	0.0050	0.2502	0.0042	0.2108	0.0077			0.1647	0.0200
	Serα (d2)	0.1214	0.0050	0.1440	0.0042	0.1274	0.0077			0.0743	0.0200
	Serα (dd)	0.3069	0.0050	0.2614	0.0042	0.3898	0.0077			0.5049	0.0200
	Serβ (enr)										
	Serβ (s)	0.5923	0.0030	0.5902	0.0036	0.4972	0.0150	0.3312	0.0280	0.4191	0.0111
	Serβ (d)	0.4071	0.0030	0.4098	0.0036	0.5028	0.0150	0.6688	0.0280	0.5809	
Pyr	Alaα (enr)	0.0500	0.0100	0.0566	0.0100	0.0616	0.0100	0.0906	0.0100	0.0745	0.0100
	Alaα (s)	0.1721	0.0060	0.2064	0.0056	0.1882	0.0101	0.1894	0.0550	0.1493	0.0305
	Alaα (d1)	0.0465	0.0060	0.0335	0.0056	0.0314	0.0101	0.0299	0.0550	0.0107	0.0305
	Alaα (d2)	0.2103	0.0060	0.1968	0.0056	0.2001	0.0101	0.1420	0.0550	0.1103	0.0305
	Alaα (dd)	0.5711	0.0060	0.5633	0.0056	0.5803	0.0101	0.6387	0.0550	0.7297	0.0305
	Alaβ (enr)	0.0360	0.0100	0.0392	0.0100	0.0861	0.0100	0.0655	0.0100	0.0532	0.0100
	Alaβ (s)	0.2459	0.0010	0.2706	0.0013	0.2443	0.0017	0.2420	0.0088	0.2382	0.0029
	Alaβ (d)	0.7541	0.0010	0.7294	0.0013	0.7557	0.0017	0.7580	0.0088	0.7618	0.0029

enr indicates the enrichment of the carbon atom from the TOCSY analysis
s indicates singlet peak, d1 indicates doublet peak, d2 the second doublet and dd the double doublet peak from the HSQC analysis

Table E.1 Continued

Precursor metabolite	Nucleus	Input Intensity H/H	Noise in spectrum H/H	Input Intensity M/H	Noise in spectrum M/H	Input Intensity M/M	Noise in spectrum M/M	Input Intensity M/L	Noise in spectrum M/L	Input Intensity L/L	Noise in spectrum L/L
Pyr	Ileγ^2 (enr)	0.0717	0.0040	0.0741	0.0100	0.0748	0.0100	0.0687	0.0100	0.0775	0.0100
	Ileγ^2 (s)	0.2519	0.0040	0.2835	0.0291	0.2415	0.0022	0.2365	0.0123	0.2117	0.0043
	Ileγ^2 (d)	0.7481	0.0040	0.7165	0.0291	0.7585	0.0022	0.7635	0.0123	0.7883	0.0043
	Leuβ (s)	0.8431	0.0160	0.8076	0.0071	0.8264	0.0071	0.8135	0.0357	0.8403	0.0155
	Leuβ (d)	0.1569	0.0160	0.1924	0.0071	0.1736	0.0071	0.1865	0.0357	0.1597	0.0155
	Leuβ (t)	0.0000	0.0160	0.0000	0.0071	0.0000	0.0071	0.0000	0.0357	0.0000	0.0155
	Leuδ^1 (enr)	0.0870	0.0100	0.0757	0.0100			0.0705	0.0100	0.0744	0.0100
	Leuδ^1 (s)	0.3336	0.0020	0.3737	0.0011	0.3569	0.0014	0.3316	0.0092	0.3137	0.0037
	Leuδ^1 (d)	0.6664	0.0020	0.6263	0.0011	0.6431	0.0014	0.6684	0.0092	0.6863	0.0037
	Leuδ^2 (s)	0.8798	0.0020	0.8885	0.0010	0.8773	0.0012	0.8805	0.0073	0.8724	0.0029
	Leuδ^2 (d)	0.1202	0.0020	0.1115	0.0010	0.1227	0.0012	0.1195	0.0073	0.1276	0.0029
	Valα (enr)	0.0796	0.0100			0.0543	0.0100	0.1034	0.0100		
	Valα (s)	0.3455	0.0100	0.3695	0.0080	0.3544	0.0265	0.3409	0.0390	0.3364	0.0113
	Valα (d1)	0.5569	0.0100	0.4989	0.0080	0.5581	0.0265	0.5629	0.0390	0.5490	0.0113
	Valα (d2)	0.0000	0.0100	0.0415	0.0080	0.0533	0.0265	0.0000	0.0390	0.0262	0.0113
	Valα (dd)	0.0976	0.0100	0.0901	0.0080	0.0342	0.0265	0.0962	0.0390	0.0884	0.0113
	Valγ^1 (enr)	0.0975	0.0100	0.1009	0.0100	0.0981	0.0100	0.0874	0.0100	0.0919	0.0100
	Valγ^1 (s)	0.2252	0.0013	0.2585	0.0016	0.2350	0.0018	0.2390	0.0090	0.2528	0.0033
	Valγ^1 (d)	0.7748	0.0013	0.7415	0.0016	0.7650	0.0018	0.7610	0.0090	0.7472	0.0033
	Valγ^2 (enr)	0.0996	0.0100	0.0960	0.0100	0.0896	0.0100	0.0874	0.0100	0.0926	0.0100
	Valγ^2 (s)	0.8652	0.0014	0.8809	0.0015	0.8820	0.0018	0.8777	0.0091	0.9026	0.0031
	Valγ^2 (d)	0.1348	0.0014	0.1193	0.0015	0.1180	0.0018	0.1223	0.0091	0.0974	0.0031

Table E.1 Continued

Precursor metabolite	Nucleus	Input Intensity H/H	Noise in spectrum H/H	Input Intensity M/H	Noise in spectrum M/H	Input Intensity M/M	Noise in spectrum M/M	Input Intensity M/L	Noise in spectrum M/L	Input Intensity L/L	Noise in spectrum L/L
PEP	Pheα (s)	0.1837	0.0100	0.1922	0.0082	0.1679	0.0119	0.1861	0.0700		
	Pheα (d1)	0.0064	0.0100	0.0537	0.0082	0.0000	0.0119	0.0671	0.0700		
	Pheα (d2)	0.1583	0.0100	0.1815	0.0082	0.1687	0.0119	0.1551	0.0700		
	Pheα (dd)	0.6516	0.0100	0.5725	0.0082	0.6634	0.0119	0.5917	0.0700		
	Pheβ (s)	0.2141	0.0120	0.2203	0.0113	0.2003	0.0110	0.1583	0.0395	0.1884	0.0227
	Pheβ (d1)	0.7444	0.0120	0.6876	0.0113	0.7539	0.0110	0.6102	0.0395	0.7163	0.0227
	Pheβ (d2)	0.0000	0.0120	0.0000	0.0113	0.0000	0.0110	0.0000	0.0395	0.0000	0.0227
	Pheβ (dd)	0.0415	0.0120	0.0921	0.0113	0.0458	0.0110	0.2315	0.0395	0.0953	0.0227
	Tyrα (s)	0.1989	0.0140	0.1743	0.0126	0.2049	0.0170			0.2020	0.0323
	Tyrα (d1)	0.0079	0.0140	0.0511	0.0126	0.0000	0.0170			0.0001	0.0323
	Tyrα (d2)	0.1111	0.0140	0.1503	0.0126	0.1083	0.0170			0.0276	0.0323
	Tyrα (dd)	0.6823	0.0140	0.6243	0.0126	0.6868	0.0170			0.7703	0.0323
	Tyrβ (s)	0.1514	0.0210	0.1856	0.0120	0.1545	0.0281	0.1069	0.0723	0.1112	0.0391
	Tyrβ (d)	0.6954	0.0210	0.7176	0.0120	0.7365	0.0281	0.3802	0.0723	0.7441	0.0391
	Tyrβ (t)	0.1532	0.0210	0.0968	0.0120	0.1090	0.0281	0.2129	0.0723	0.1447	0.0391
ACoA	Leuα (enr)	0.0606	0.0100			0.0948	0.0000	0.1026	0.0100	0.1010	0.0100
	Leuα (s)	0.2252	0.0050	0.2514	0.0096	0.2378	0.0088	0.2853	0.0414	0.2370	0.0127
	Leuα (d1)	0.7265	0.0050	0.7059	0.0096	0.7251	0.0088	0.6848	0.0414	0.6800	0.0127
	Leuα (d2)	0.0169	0.0050	0.0199	0.0096	0.0320	0.0088	0.0162	0.0414	0.0000	0.0127
	Hisβ (s)	0.3264	0.0070	0.3811	0.0068	0.3318	0.0106	0.3779	0.0585	0.3249	0.0194
	Hisβ (d1)	0.0541	0.0070	0.0000	0.0068	0.0000	0.0106	0.0000	0.0585	0.0000	0.0194
	Hisβ (d2)	0.5349	0.0070	0.5232	0.0068	0.5778	0.0106	0.6221	0.0585	0.5629	0.0194
	Hisβ (dd)	0.0846	0.0070	0.0957	0.0068	0.0903	0.0106	0.0000	0.0585	0.1122	0.0194

Table E.1 Continued

Precursor metabolite	Nucleus	Input Intensity H/H	Noise in spectrum H/H	Input Intensity M/H	Noise in spectrum M/H	Input Intensity M/M	Noise in spectrum M/M	Input Intensity M/L	Noise in spectrum M/L	Input Intensity L/L	Noise in spectrum L/L
P5P	Hisδ^2(s)	0.6744	0.0500	0.6782	0.0332	0.6917	0.0538			0.3618	0.2794
	Hisδ^2(d)	0.3256	0.0500	0.3218	0.0332	0.3083	0.0538			0.6382	0.2794
E4P	Pheδ (enr)							0.1056	0.0100		
	Tyrδ (s)	0.2667	0.0500	0.3402	0.0734	0.3313	0.1143			0.2690	0.2496
	Tyrδ (d)	0.7333	0.0500	0.6598	0.0734	0.6687	0.1143			0.7330	0.2496
	Tyrδ (t)	0.0000	0.0500	0.0000	0.0734	0.0000	0.1143			0.0000	0.2496
Pyr+OAA	Lysα (enr)	0.0798	0.0100	0.0573	0.0100	0.0682	0.0100	0.0906	0.0100	0.0733	0.0100
	Lysβ (s)	0.3175	0.0130	0.2571	0.0138	0.2348	0.0130	0.1258	0.0456	0.3143	0.0346
	Lysβ (d)	0.6825	0.0130	0.6284	0.0138	0.6597	0.0130	0.5570	0.0456	0.6857	0.0346
	Lysβ (t)	0.0000	0.0130	0.1145	0.0138	0.1055	0.0130	0.3172	0.0456	0.0000	0.0346
	Lysδ (s)	0.2589	0.0060	0.2625	0.0172	0.2502	0.0080	0.2454	0.0291	0.3248	0.0144
	Lysδ (d)	0.6418	0.0060	0.6312	0.0172	0.6418	0.0080	0.5973	0.0291	0.6752	0.0144
	Lysδ (t)	0.0993	0.0060	0.1063	0.0172	0.1080	0.0080	0.1573	0.0291	0.0000	0.0144
	Lysε (enr)	0.0538	0.0100	0.0530	0.0100	0.0569	0.0100	0.0356	0.0100	0.0543	0.0100
	Lysε (s)	0.3329	0.0030	0.3395	0.0042	0.3181	0.0047	0.3457	0.0274	0.3435	0.0100
	Lysε (d)	0.6671	0.0030	0.6605	0.0042	0.6819	0.0047	0.6543	0.0274	0.6565	0.0100

Table E.1 Continued

Precursor metabolite	Nucleus	Input Intensity H/H	Noise in spectrum H/H	Input Intensity M/H	Noise in spectrum M/H	Input Intensity M/M	Noise in spectrum M/M	Input Intensity M/L	Noise in spectrum M/L	Input Intensity L/L	Noise in spectrum L/L
Glu	Argβ (s)	0.8274	0.0400	0.7945	0.0480	0.8298	0.0526			0.7352	0.0879
	Argβ (d)	0.1726	0.0400	0.2055	0.0480	0.1702	0.0526			0.2648	0.0879
	Argβ (t)	0.0000	0.0400	0.0000	0.0480	0.0000	0.0526			0.0000	0.0879
	Argγ (s)	0.6045	0.0530								
	Argγ (d)	0.3955	0.0530								
	Argγ (t)	0.0000	0.0530								
	Argδ (s)	0.5898	0.0060	0.6166	0.0084	0.6629	0.0122	0.5516	0.0480	0.5610	0.0148
	Argδ (t)	0.4102	0.0060	0.3834	0.0084	0.3371	0.0122	0.4484	0.0480	0.4390	0.0148
	Gluβ (s)			0.7632	0.0130	0.8210	0.0173	0.8203	0.0535	0.8149	0.0259
	Gluβ (d)			0.2368	0.0130	0.1790	0.0173	0.1797	0.0535	0.1851	0.0259
	Gluβ (t)			0.0000	0.0130	0.0000	0.0173	0.0000	0.0535	0.0000	0.0259
	Gluγ (s)	0.5937	0.0040	0.5953	0.0048	0.6944	0.0089	0.6028	0.0199		
	Gluγ (d1)	0.0345	0.0040	0.0318	0.0048	0.0158	0.0089	0.0000	0.0199		
	Gluγ (d2)	0.3718	0.0040	0.3729	0.0048	0.2898	0.0089	0.3972	0.0199		
	Gluγ (dd)	0.0000	0.0040	0.0000	0.0048	0.0000	0.0089	0.0000	0.0199		
	Proγ (s)	0.5623	0.0300	0.5641	0.0180	0.6393	0.0290	0.4950	0.1212	0.5364	0.0475
	Proγ (d)	0.4377	0.0300	0.4359	0.0180	0.3607	0.0290	0.5050	0.1212	0.4636	0.0475
	Proγ (t)	0.0000	0.0300	0.0000	0.0180	0.0000	0.0290	0.0000	0.1212	0.0000	0.0475
	Proδ (s)	0.5663	0.0220	0.5805	0.0172	0.6414	0.0361	0.4678	0.1384	0.4768	0.0740
OAA	Aspα (s)	0.4559	0.0100	0.4511	0.0062	0.4326	0.0100	0.5467	0.0460	0.4300	0.0166
	Aspα (d1)	0.1463	0.0100	0.1312	0.0062	0.0872	0.0100	0.1474	0.0460	0.1428	0.0166
	Aspα (d2)	0.2195	0.0100	0.2555	0.0062	0.2832	0.0100	0.2124	0.0460	0.1734	0.0166
	Aspα (dd)	0.1783	0.0100	0.1622	0.0062	0.1970	0.0100	0.0934	0.0460	0.2538	0.0166

Table E.1 Continued

Precursor metabolite	Nucleus	Input Intensity H/H	Noise in spectrum H/H	Input Intensity M/H	Noise in spectrum M/H	Input Intensity M/M	Noise in spectrum M/M	Input Intensity M/L	Noise in spectrum M/L	Input Intensity L/L	Noise in spectrum L/L
OAA	Aspβ (s)	0.4446	0.0100	0.4240	0.0031	0.4059	0.0100	0.4362	0.0173	0.4822	0.0074
	Aspβ (d1)	0.2897	0.0100	0.3230	0.0031	0.3434	0.0100	0.2585	0.0173	0.2569	0.0074
	Aspβ (d2)	0.1129	0.0100	0.1086	0.0031	0.0818	0.0100	0.1396	0.0173	0.1493	0.0074
	Aspβ (dd)	0.1528	0.0100	0.1444	0.0031	0.1690	0.0100	0.1657	0.0173	0.1116	0.0074
	Ileα (s)	0.5827	0.0160	0.6135	0.0100	0.5956	0.0130	0.6076	0.1128	0.6058	0.0299
	Ileα (d1)	0.3158	0.0160	0.3005	0.0100	0.2870	0.0130	0.3924	0.1128	0.2974	0.0299
	Ileα (d2)	0.1015	0.0160	0.0860	0.0100	0.0598	0.0130	0.0000	0.1128	0.0968	0.0299
	Ileα (dd)	0.0000	0.0160	0.0000	0.0100	0.0576	0.0130	0.0000	0.1128	0.0000	0.0299
	Ileγ^1(s)	0.6460	0.0450	0.2835	0.0291	0.6603	0.0315	0.5845	0.1000	0.6713	0.0717
	Ileγ^1(d)	0.3540	0.0450	0.7165	0.0291	0.3397	0.0315	0.4155	0.1000	0.3287	0.0717
	Ileγ^1(t)	0.0000	0.0450	0.0000	0.0291	0.0000	0.0315	0.0000	0.1000	0.0000	0.0717
	Ileδ (s)	0.6885	0.0050	0.7292	0.0021	0.7317	0.0065	0.6960	0.0123	0.7132	0.0125
	Ileδ (d)	0.3115	0.0050	0.2708	0.0021	0.2683	0.0065	0.3040	0.0123	0.2868	0.0125
	Metγ (s)	0.6965	0.0300	0.7010	0.0312	0.7297	0.0332	0.6415	0.1580	0.7075	0.0583
	Metγ (d)	0.3035	0.0300	0.2990	0.0312	0.2703	0.0332	0.3585	0.1580	0.2925	0.0583
	Lysγ (s)	0.6625	0.0320	0.6591	0.0172	0.6546	0.0232	0.5638	0.0844	0.6271	0.0553
	Lysγ (d)	0.3375	0.0320	0.3409	0.0172	0.3454	0.0232	0.4362	0.0844	0.3729	0.0553
	Lysγ (t)	0.0000	0.0320	0.0000	0.0172	0.0000	0.0232	0.0000	0.0844	0.0000	0.0553

Table E.1 Continued

Precursor metabolite	Nucleus	Input Intensity H/H	Noise in spectrum H/H	Input Intensity M/H	Noise in spectrum M/H	Input Intensity M/M	Noise in spectrum M/M	Input Intensity M/L	Noise in spectrum M/L	Input Intensity L/L	Noise in spectrum L/L
OAA	Thrα (s)	0.3483	0.0130	0.4068	0.0010	0.4034	0.0500			0.4448	0.0361
	Thrα (d1)	0.1483	0.0130	0.1923	0.0010	0.1218	0.0500			0.2295	0.0361
	Thrα (d2)	0.2469	0.0130	0.1988	0.0010	0.2251	0.0500			0.0704	0.0361
	Thrα (dd)	0.2565	0.0130	0.2021	0.0010	0.2497	0.0500			0.2553	0.0361
	Thrβ (enr)	0.0247	0.0100	0.0331	0.0100	0.0496	0.0100	0.0378	0.0100	0.0488	0.0100
	Thrβ (s)	0.5104	0.1150	0.4859	0.0810	0.5115	0.1040			0.3071	0.1877
	Thrβ (d)	0.4896	0.1150	0.5141	0.0810	0.4885	0.1040			0.6929	0.1877
	Thrβ (t)	0.0000	0.1150	0.0000	0.0810	0.0000	0.1040			0.0000	0.1877
	Thrγ^2 (enr)	0.0208	0.0100	0.0669	0.0100					0.0543	0.0100
	Thrγ^2 (s)	0.6837	0.0040	0.7055	0.0044	0.7094	0.0064	0.6776	0.0290		
	Thrγ^2 (d)	0.3163	0.0040	0.2945	0.0044	0.2906	0.0064	0.3224	0.0290		

Appendix F: Metabolic flux maps for H/H, M/H, M/L and L/L temperature treatments

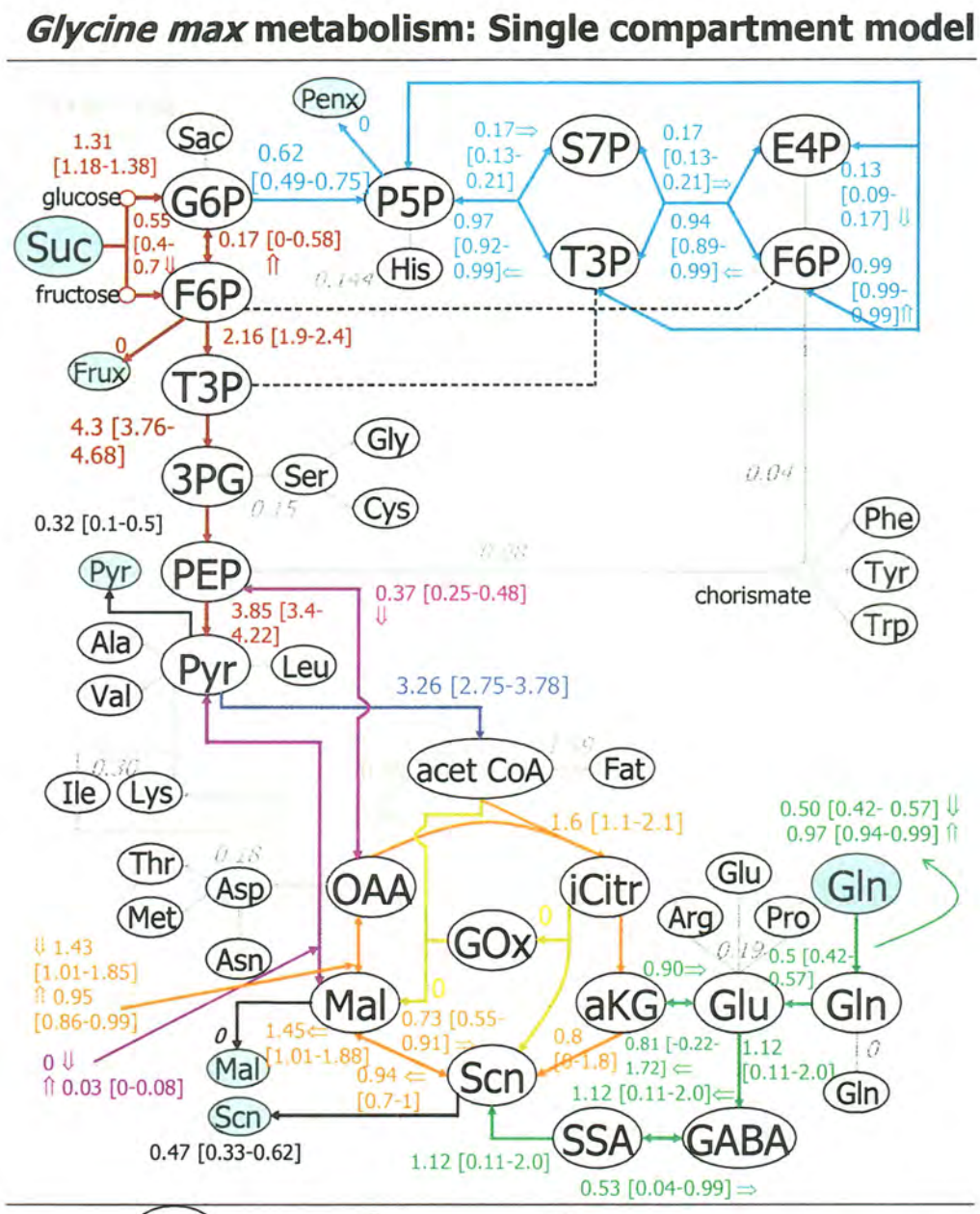


Figure F.1 Metabolic flux map for soybean embryos grown under H/H (Flux values are standardized with respect to 20 times glycine production per day)

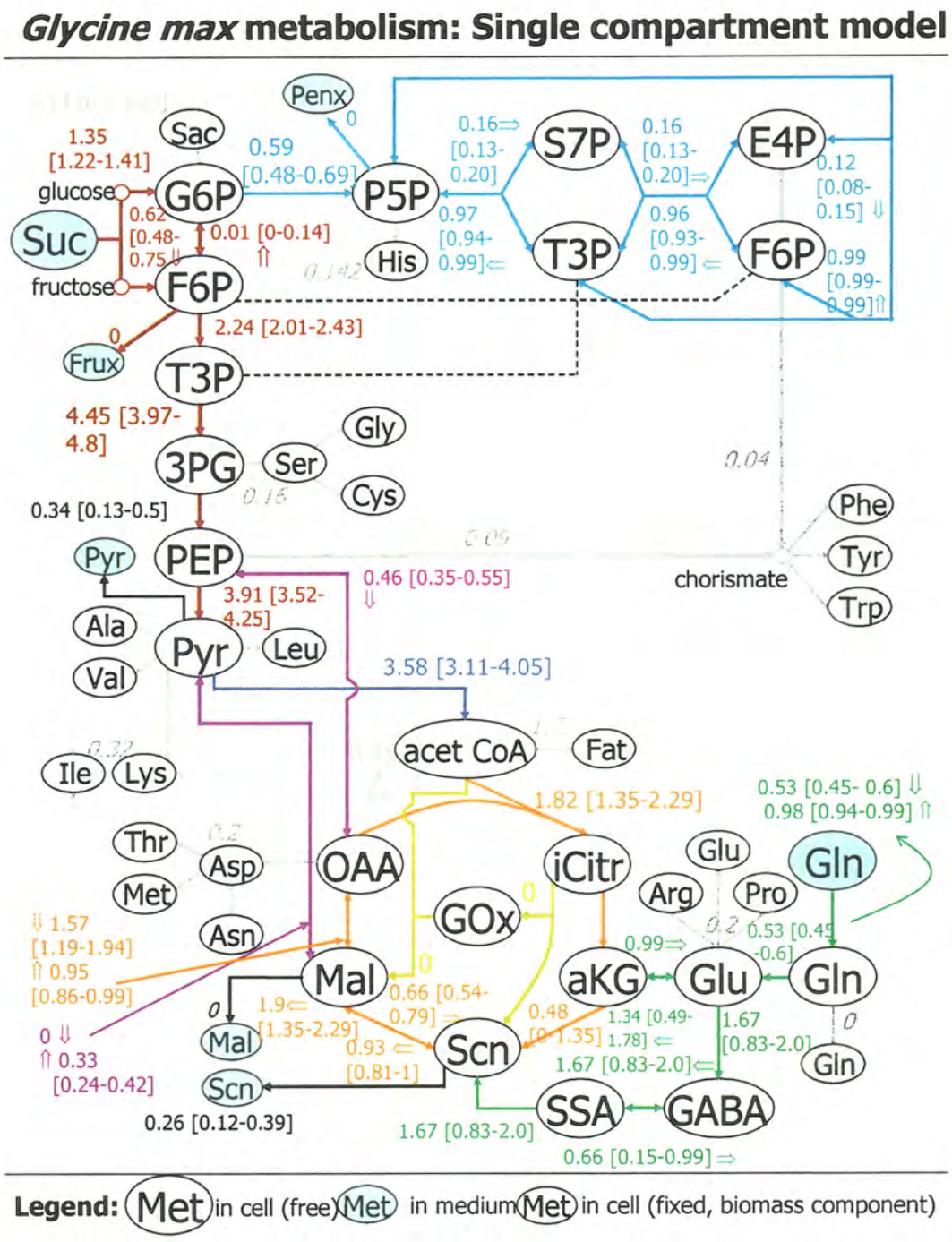


Figure F.2 Metabolic flux map for soybean embryos grown under M/H (Flux values are standardized with respect to 20 times glycine production per day)

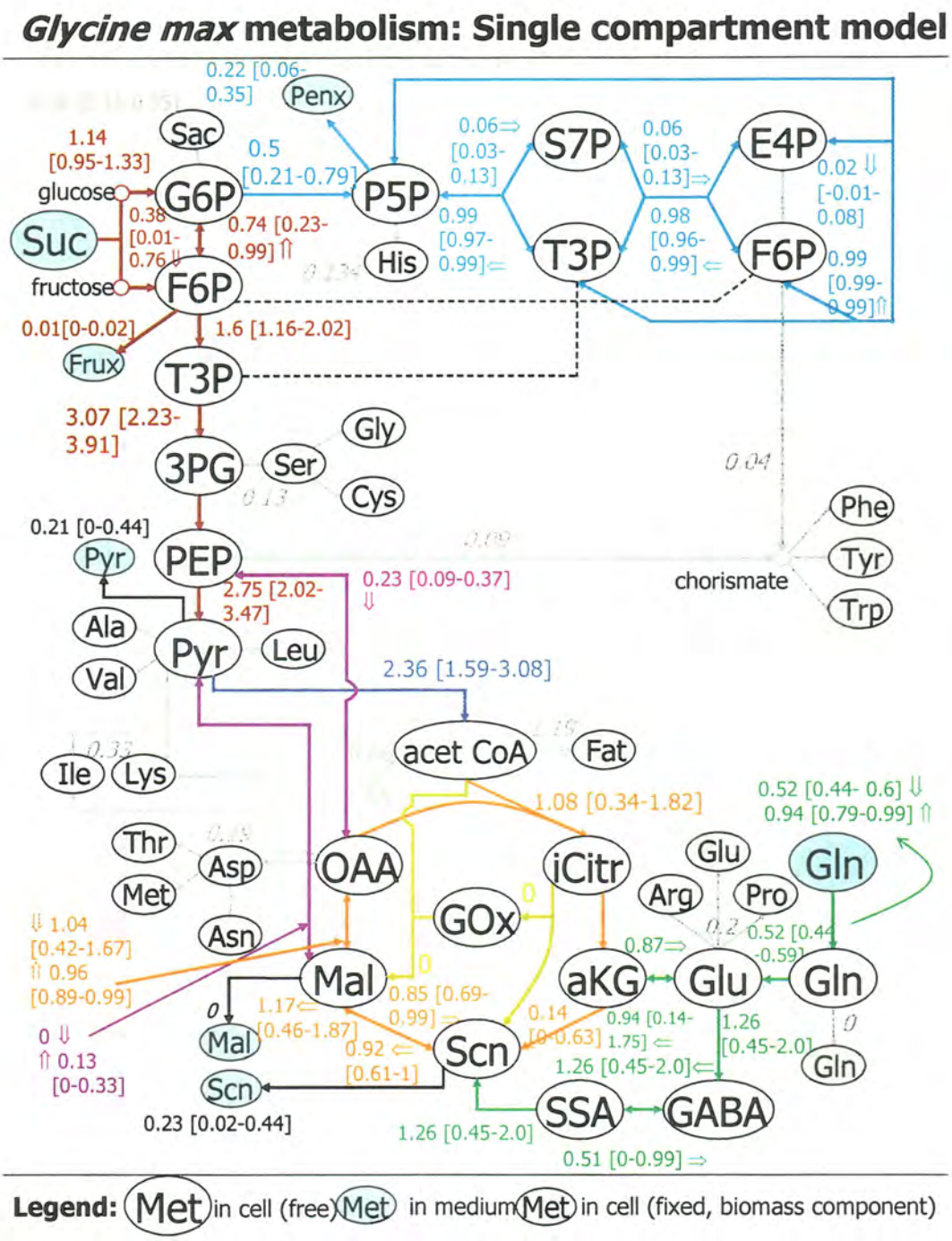


Figure F.3 Metabolic flux map for soybean embryos grown under M/L (Flux values are standardized with respect to 20 times glycine production per day)

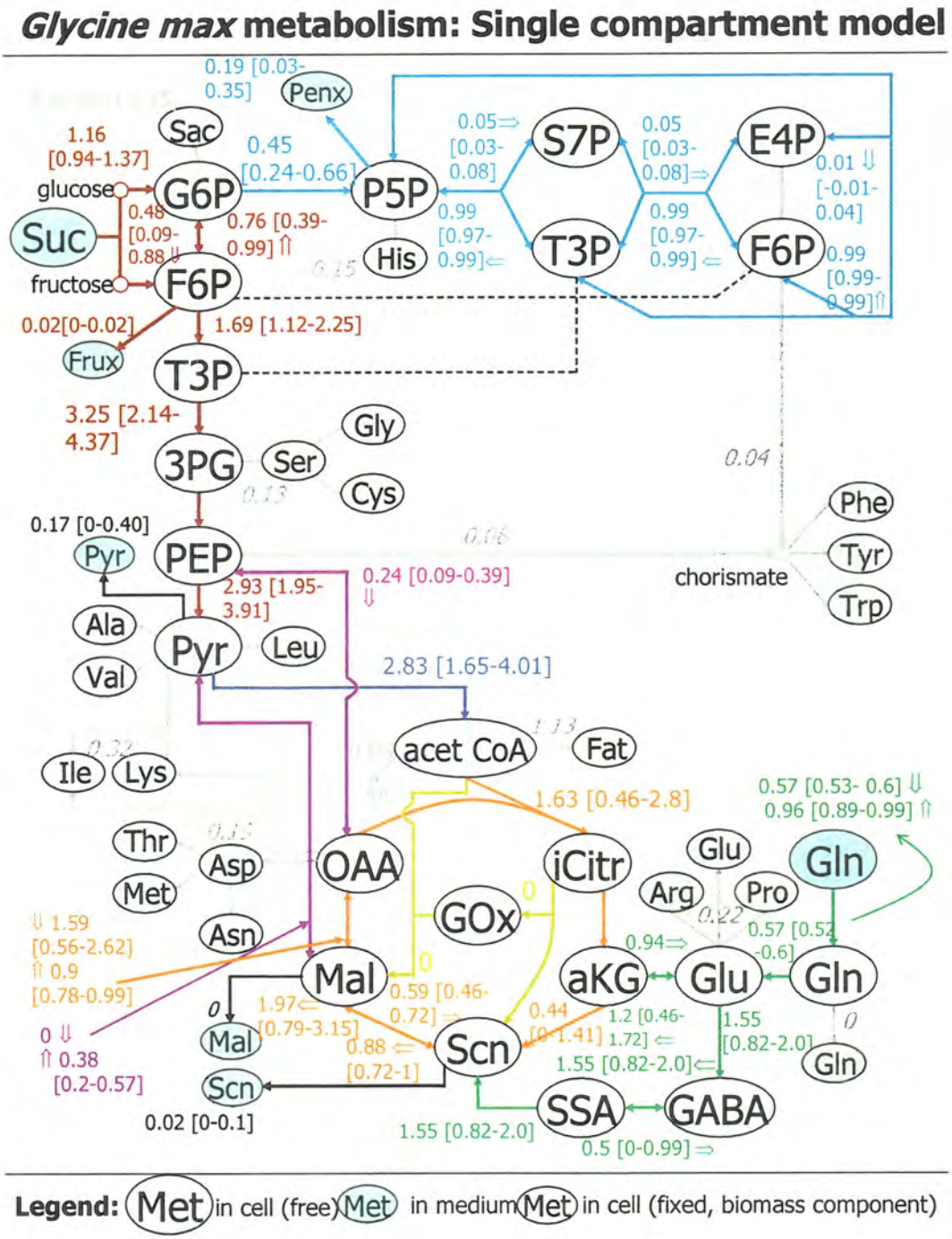


Figure F.4 Metabolic flux map for soybean embryos grown under L/L (Flux values are standardized with respect to 20 times glycine production per day)

BIBLIOGRAPHY

1. Myers, D.J., *Industrial applications for soy protein and potential for increased utilization.* Cereal foods world, 1993. **38**: p. 355-360.
2. Kinney, A., J, *Plants as industrial chemical factories-new oils from genetically engineered soybeans.* Fett/Lipid, 1998. **100**: p. 173-176.
3. Maestri, D., M, et al., *Seed composition of soybean cultivars evaluated in different environmental regions.* Journal of the science of food and agriculture, 1998. 77: p. 494-498.
4. Hayati, R., D. Egli, B, and S. Rafts-Bradner, J, *Independence of nitrogen supply and seed growth in soybeans: studies using an in vitro culture system.* Journal of experimental botany, 1996. **47**: p. 33-40.
5. Stitt, M. and A.R. Fernie, *From measurements of metabolites to metabolomics: an 'on the fly' perspective illustrated by recent studies of carbon-nitrogen interactions.* Current Opinion in Biotechnology, 2003. **14**: p. 136-144.
6. Szyperski, T., *Biosynthetically directed fractional ^{13}C-labeling of proteinogenic amino acids. An efficient analytical tool to investigate intermediary metabolism.* European Journal of Biochemistry, 1995. **232**: p. 433-448.
7. Szyperski, T., et al., *Bioreaction network topology and metabolic flux ratio analysis by biosynthetic fractional ^{13}C-labeling and two-dimensional NMR spectroscopy.* 1999.
8. Wiechert, W., et al., *A universal framework for ^{13}C metabolic flux analysis.* Metabolic engineering, 2001. **3**: p. 265-283.
9. Sriram, G. and J.V. Shanks, *Metabolic flux analysis using carbon bond labeling experiments and bondomer balancing: analytical solutions and numerical innovations.* Metabolic engineering, 2003.
10. Sweetlove, L.J., R.L. Last, and A.R. Fernie, *Predictive metabolic engineering: A goal for systems biology.* Plant Physiology, 2003. **132**: p. 420-425.
11. Christensen, B. and J. Nielsen, *Isotopomer analysis using GC-MS.* Metabolic engineering, 1999. **1**: p. 282-290.

12. Szyperski, T., *^{13}C-NMR, MS and metabolic flux balancing in biotechnology research.* Quarterly reviews of Biophysics, 1998. **31**(1): p. 41-106.

13. Ratcliffe, R.G. and Y. Shachar-Hill, *Probing plant metabolism with NMR.* Annual Review of Plant Physiology and Plant Molecular Biology, 2001. **52**: p. 499-526.

14. Roscher, A., N.J. Kruger, and R.G. Ratcliffe, *Strategies from metabolic flux analysis in plants using isotope labeling.* Journal of Biotechnology, 2000. **77**: p. 81-102.

15. Dieuaide-Noubhani, M., et al., *Quantification of compartmented metabolic fluxes in maize root tips using isotope distribution from ^{13}C or ^{14}C-labeled glucose.* The Journal of biological chemistry, 1995. **270**: p. 13147-13159.

16. Eichiger, D., et al., *Quantitative assessment of metabolic flux by ^{13}C NMR analysis. Biosynthesis of anthraquinones in Rubia tinctorum.* Journal of american chemical society, 1999. **121**: p. 7469-7475.

17. Glawschnig, E., et al., *Retrobiosynthetic nuclear magnetic resonance analysis of amino acid biosynthesis and intermediary metabolism. Metabolic flux in developing maize kernels.* Plant Physiology, 2001. **125**: p. 1178-1186.

18. Glawschnig, E., et al., *Starch biosynthesis and intermediary metabolism in maize kernels. Quantitative analysis of metabolite flux by nuclear magnetic resonance.* Plant Physiology, 2002. **130**: p. 1717-1727.

19. Rijhwani, S.K., C.-H. Ho, and J.V. Shanks, *In vivo ^{31}P and multilabel ^{13}C NMR measurements for evaluation of plant metabolic pathways.* Metabolic engineering, 1999. **1**: p. 12-25.

20. Nielsen, J., *Metabolic engineering: Techniques for analysis of targets for genetic manipulations.* Biotechnology and Bioengineering, 1998. **58**: p. 125-132.

21. Varma, A. and B.O. Palsson, *Metabolic flux balancing: Basic concepts, scientific and practical use.* Bio/Technology, 1994. **12**: p. 994-998.

22. Stephanopoulos, G., J. Nielsen, and A.A. Aristidou, *Metabolic engineering: Principles and methodologies.* First ed. 1998: Elsevier Science & Technology Books.

23. Schmidt, K., et al., *^{13}C tracer experiments and metabolite balancing for metabolic flux analysis: Comparing two approaches.* Biotechnology and Bioengineering, 1998. **58**: p. 254-257.

24. Wiechert, W., *Minireview ^{13}C Metabolic flux analysis.* Metabolic engineering, 2001. **3**: p. 195-206.

25. Marx, A., et al., *Determination of the fluxes in the central metabolism of Corynebacterium glutamicum by nuclear magnetic resonance spectroscopy combined with metabolite balancing.* Biotechnology and Bioengineering, 1996. **49**: p. 111-129.

26. Klapa, M.I., et al., *Metabolite and isotopomer balancing in the analysis of metabolic cycles: I. Theory.* Biotechnology and Bioengineering, 1999. **62**(4): p. 375-391.

27. Follstad, B.D. and G. Stephanopoulos, *Effect of reversible reactions on isotope label redistribution. Analysis of the pentose phosphate pathway.* European Journal of Biochemistry, 1998. **252**: p. 360-371.

28. Schmidt, K., et al., *Modeling isotopomer distributions in biochemical networks using isotopomer mapping matrices.* Biotechnology and Bioengineering, 1997. **55**(6): p. 831-840.

29. Zupke, C. and G. Stephanopoulos, *Modeling of isotope distributions and intracellular fluxes in metabolic networks using atom mapping matrices.* Biotechnology Progress, 1994. **10**: p. 489-498.

30. Schmidt, K., J. Nielsen, and J. Villadsen, *Quantitative analysis of metabolite fluxes in Escherichia coli, using two-dimensional NMR spectroscopy and complete isotopomer models.* Journal of Biotechnology, 1999. **71**: p. 175-190.

31. Sriram, G. and J.V. Shanks, *A mathematical model for carbon bond labeling experiments: Analytical solutions and sensitivity analysis for the effect of reaction reversibilities on estimated fluxes*, in *Chemical Engineering.* 2002, Kansas State University: Kansas City. p. 45-54.

32. van Winden, W., J. Heijnen, and P. Verheijen, *Cumulative bondomers: a new concept in flux analysis from 2D ^{13}C, ^{1}H COSY NMR data.* Biotechnology and Bioengineering, 2002. **80**: p. 731-745.

33. Wiechert, W. and A.A. De Graaf, *Bidirectional reaction steps in metabolic networks: I Modeling and simulation of carbon isotope labeling experiments.* Biotechnology and Bioengineering, 1997a. **55**: p. 101-117.

34. Schmidt, K., et al., *Quantification of intracellular metabolic fluxes from fractional enrichment ^{13}C-^{13}C coupling constraints on the isotopomer distribution in labeled biomass components.* Metabolic engineering, 1999. **1**: p. 166-179.

35. Park, S.M., et al., *Metabolite and isotopomer balancing in the analysis of metabolic cycles: II. Applications.* Biotechnology and Bioengineering, 1999. **62**(4): p. 392-401.

36. Mollney, M., et al., *Bidirectional reaction steps in metabolic networks: IV. Optimal design of isotopomer labeling experiments.* Biotechnology and Bioengineering, 1999. **66**: p. 86-103.

37. Wiechert, W., et al., *Bidirectional reaction steps in metabolic networks:III Explicit solution and analysis of isotopomer labeling systems.* Biotechnology and Bioengineering, 1999. **66**: p. 69-85.

38. Sriram, G., et al., *Quantification of metabolic fluxes in developing soybean embryos using biosynthetically directed fractional ^{13}C labeling, 2-D [^{13}C, ^{1}H] NMR and rigorous isotopomer balancing.* Plant Physiology, 2003.

39. Sontagg, K., et al., *Flux partitioning in the split pathway of lysine synthesis in Corynebacterium glutamicum. Quantification by ^{13}C and ^{1}H- NMR spectroscopy.* European Journal of Biochemistry, 1993. **213**: p. 1325-1331.

40. Dauner, M., et al., *Intracellular carbon fluxes in riboflavin-producing Bacillus subtilis during growth on two-carbon substrate mixtures.* Applied Environmental Microbiology, 2002. **68**: p. 1760-1771.

41. Dauner, M., J.E. Bailey, and U. Sauer, *Metabolic flux analysis with a comprehensice isotopomer model in Bacillus subtilis.* Biotechnology and Bioengineering, 2000. **76**: p. 144-156.

42. Emmerling, M., et al., *Metabolic flux responses to pyruvate kinase knockout in Escherichia coli.* Journal of Biotechnology, 2002. **184**: p. 152-164.

43. Dornbos, D., L Jr. and M. Mc Donald, B Jr., *Mass and composition of developing soybean seeds at five reproductive stages.* Crop Science, 1986. **26**: p. 624-630.

44. Guo, C. and D. Oosterhuis, M, *Pinitol occurence in soybean plants as affected by temperature and plant growth regulators.* Journal of experimental botany, 1995. **46**: p. 249-253.

45. Saravitz, C., H and D. Raper, C, Jr., *Responses to sucrose and glutamine by soybean embryos grown in vitro.* Physiologia plantarum, 1995. **93**: p. 799-805.

46. Obendorf, R.L., G.T. Rytko, and M.C. Byrne, *Soybean seed growth and maturation by in vitro pod culture.* Annals of Botany, 1983. **51**: p. 217-227.

47. Rivera, O.G., G. Sriram, and J.V. Shanks. *Determination of the biomass composition of hairy roots of Catharantus roseus.* in *32nd Annual biochemical engineering symposium.* 2002. Ames, Iowa.

48. Chollet, R., J. Vidal, and M.H. O'Leary, *Phosphoenolpyruvate carboxylase: A ubiquitous, highly regulated enzyme in plants.* Annual Review of Plant Physiology and Plant Molecular Biology, 1996. **47**: p. 273-98.

49. Casati, P., et al., *Malate metabolism by NADP-malic enzyme in plant defense.* Photosynthesis research, 1999. **61**: p. 99-105.

50. Bennett, J.M., et al., *Respiration and carbon assimilate use*, in *Physiology and Determination of Crop Yield.* 1994. p. 237.

51. Nakamura, A., et al., *Analysis of structural components and molecular construction of soybean soluble polysaccharides by stepwise enzymatic degradation.* Bioscience, Biotechnology and Biochemsitry, 2001. **65**: p. 2249-58.

52. Shanks, J.V., *In Situ NMR systems*, in *NMR in microbiology: Theory and applications*, J.N. Barbotin and J.C. Portais, Editors. 2000, Horizon scientific press. p. 49-75.

53. Kumar, S., et al., *Metabolic fate of glutamate and evaluation of flux through the 4-aminobutyrate (GABA) shunt in Aspergillus niger.* Biotechnology and Bioengineering, 2000. **67**: p. 575-584.

54. Stephanopoulos, G. and J.J. Vallino, *Network rigidity and metabolic engineering in metabolite overproduction.* Science, New Series, 1991. **252**(5013): p. 1675-1681.

55. Heldt, H.-W., *Plant biochemistry and molecular biology*, ed. H.-W. Heldt. 1997: Oxford press.

ACKOWLEDGEMENTS

I thank Dr. Jackie Shanks for her invaluable guidance and giving me an opportunity to pursue my MS degree. I thank National Science Foundation and the Plant Facility at Iowa State University for funding this work. My sincere thanks to Ganesh, for mentoring me throughout and patiently answering all my questions. For serving on my committee and their valuable suggestions on my thesis, I thank Dr. Mark Westgate and Dr. Ramon Gonzalez. I have to mention Ruilian, Dr. J. Peterson and Dr. Bruce Fulton who helped me with my experimentation.

For the guidance and encouragement through my thesis writing, I am grateful to my friends Mothi, Chandrika, Rahul and Sudershan. I would also like to thank my friends Rajee, Vasu, Kamal, Dinesh, Murali, Shantha, and Anil for making my stay in Ames a memorable one.

I thank my parents, my brother Vinay and my childhood friend Sangita for their support and affection from miles away. And finally, I thank Kishore for his love and support and always being there for me.